VOL.
4
1958

THE PHYSICS AND CHEMISTRY OF SOLIDS

AN INTERNATIONAL JOURNAL

VOLUME 4 1958

Editor-in-Chief HARVEY BROOKS *Cambridge, Mass.*

Associate Editors
H. B. G. CASIMIR, *Eindhoven*
G. J. DIENES, *Brookhaven*
JACQUES FRIEDEL, *Paris*
L. D. LANDAU, *Moscow*
E. M. LIFSHITZ, *Moscow*

Editorial Advisory Board

U.S.A.: L. Apker, Schenectady · J. Bardeen, Illinois · D. S. Billington, Oak Ridge · Leo Brewer, Berkeley, Calif. · N. Cabrera, Virginia · I. Estermann, Washington, D.C. · J. C. Fisher, Schenectady · Conyers Herring, New Jersey · A. R. von Hippel, Cambridge, Mass. · J. H. Hollomon, Schenectady · W. V. Houston, Houston · H. B. Huntington, Troy · John G. Kirkwood, Yale · K. Lark-Horowitz, Purdue · A. W. Lawson, Chicago · B. T. Matthias, New Jersey · Joseph E. Mayer, Chicago · David Pines, Princeton R. Smoluchowski, Pittsburgh · D. Turnbull, Schenectady · J. H. Van Vleck, Harvard · B. E. Warren, Cambridge, Mass. · Eugene P. Wigner, Princeton · B. H. Zimm, Schenectady.

United Kingdom: A. Charlesby, Shrivenham, Berks · W. J. Dunning and F. C. Frank, Bristol · H. K. Henisch, Reading · H. Jones, London · J. W. Mitchell, Bristol · A. B. Pippard, Cambridge · M. H. L. Pryce, Bristol F. C. Tompkins, London.

France: E. F. Bertaut, Grenoble · A. Guinier, Paris · Louis Néel and Louis Weil, Grenoble.

Germany: M. Born, Bad Pyrmont · Ulrich Dehlinger, Stuttgart · R. Hilsch and G. Leibfried, Göttingen E. Mollwo, Erlangen · H. Pick, Stuttgart · M. von Laue, Berlin · H. Welker, Erlangen.

Italy: F. G. Fumi, Pavia.

Japan: Toshinosuke Muto, Tokyo.

The Netherlands: A. Michels and G. W. Rathenau, Amsterdam. *Sweden:* J. O. Linde, Stockholm · Per-Olov Löwdin, Uppsala.

Switzerland: G. Busch and P. Scherrer, Zürich. *U.S.S.R.:* G. S. Zhdanov, Moscow.

PERGAMON PRESS NEW YORK · LONDON · PARIS · LOS ANGELES

NOTES FOR CONTRIBUTORS

I. GENERAL

1. Submission of a paper to the *Journal of the Physics and Chemistry of Solids* will be taken to imply that it represents original research not previously published (except in the form of an abstract or preliminary report), that it is not being considered for publication elsewhere, and that if accepted, it will not be published elsewhere in the same form, in any language, without the consent of the editor-in-chief. It should deal with original research work in the field of the physics and chemistry of solids.

2. Papers should be submitted to the appropriate regional editor (all English-language papers to be sent to the U.S. editor).

3. Papers will be published as quickly as possible after acceptance, and, subject to space being available, should appear in the following issue, if this is due for publication not earlier than three months after the acceptance date. Short communications under the heading of "Letters to the Editors" will receive priority for publication, and will be published in the issue following receipt, if accepted not later than the beginning of the month preceding publication.

4. Fifty free reprints of each paper are supplied. Additional copies can be obtained at a reasonable cost if ordered when proofs are returned. A reprint order form will accompany first proofs.

II. SCRIPT REQUIREMENTS

1. Papers submitted should be concise and written in a readily understandable style. Scripts should be typed and double spaced and submitted in duplicate to facilitate refereeing.

It will be appreciated if authors clearly indicate any special characters used. An abstract, not exceeding 200 words, should be provided in the language of the paper. French and German papers should be submitted with English abstract and titles, but if this is not possible the abstract will be translated by the publishers. To conserve space, authors are requested to mark less important parts of the paper, such as details of experimental technique, methods, mathematical derivations, etc. for printing in small type. The technical description of methods should be given in detail only when such methods are new. Authors will receive proofs for correction when their papers are first set; page proofs will be sent only when the amount of alteration makes it advisable.

2. Illustrations should not be included in the typescript of the paper, and legends should be typed on a separate sheet. Line drawings which require redrawing should include all relevant details and clear instructions for the draughtsman. If figures are already well drawn it may be possible to reproduce them direct from the originals, or from good photo-prints if these can be provided. It is not possible to reproduce from prints with weak lines. Illustrations for reproduction should normally be about twice the final size required. Photographs should only be included where they are essential.

Tables and figures should be so constructed as to be intelligible without reference to the text. Every table and column should be provided with an explanatory heading. Units of measure must always be clearly indicated. The same data should not be published in both tables and figures. The following standard symbols should be used on line drawings since they are easily available to the printers: ○, ●, +, ×, □, ■, △, ▲, ◇, ◆, ▽, ▼.

4. References are indicated in the text by superior numbers in parentheses, and the full reference should be given in a list at the end of the paper in the following form:

1. Hill R., *J. Iron St. Inst.* **158,** 177 (1948).
2. Pearson C., *The Extrusion of Metals*. Chapman and Hall, London (1944).

Abbreviations of journal titles should follow those given for *Physics Abstracts*. It is particularly requested that authors' initials, and appropriate volume and page numbers, should be given in every case.

Footnotes, as distinct from literature references should be indicated by the following symbols—*, †, ‡, ‡‡, commencing anew on each page; they should *not* be included in the numbered reference system.

5. Due to the international character of the journal no rigid rules concerning notation and spelling will be observed, but each paper should be consistent within itself as to symbols and units.

VOL.
4
1958

VOL
4
195

CONTENTS

NUMBERS 1/2

E. M. LIFSHITZ and A. M. KOSEVICH: Theory of the Shubnikov-de Haas effect 1

E. M. LIFSHITZ: Quantum theory of the electrical conductivity of metals in a magnetic field . 11

P. N. ARGYRES: Quantum theory of longitudinal magnetoresistance 19

GERHARD BARSCH: Wahrscheinlichkeitsverteilung und Korrelation von Schubspannungen im Kristall 27

ERNST FELDTKELLER: Zum Energieumsatz bei röntgenphotochemischen Prozessen in KBr-Kristallen 50

HOWARD REISS and C. S. FULLER: The effect of ion pair and ion triplet formation on the solubility of lithium in germanium-effect of gallium and zinc 58

J. P. MAITA: Ion pairing in silicon 68

P. G. DE GENNES et J. FRIEDEL: Anomalies de résistivité dans certains métaux magnétiques . . 71

H. J. KRUSEMEYER and D. G. THOMAS: Adsorption and charge transfer on semiconductor surfaces. 78

J. T. LAW: The interaction of oxygen with clean silicon surfaces 91

RICHARD F. WALLIS: Theory of cyclotron-resonance absorption by conduction electrons in indium antimonide 101

E. DANIEL et J. FRIEDEL: Sur l'annihilation du positon dans les métaux 111

H. J. JURETSCHKE and R. STEINITZ: Hall effect and electrical conductivity of transition-metal diborides 118

D. R. WINDER and CHARLES S. SMITH: Single-crystal elastic constants of indium 128

R. J. WEISS and K. J. TAUER: Thermodynamics and magnetic structures of the allotropic modifications of manganese 135

W. P. DOYLE: Absorption spectra of solids and chemical bonding—I Arsenic, antimony and bismuth trioxides 144

H. J. HROSTOWSKI and R. H. KAISER: Infrared spectra of group III acceptors in silicon 148

Letters to the Editors:

J. H. WERNICK, S. GELLER and K. E. BENSON: New semiconductors 154

H. J. HROSTOWSKI and C. S. FULLER: Extension of infrared spectra of III–V compounds by lithium diffusion. 155

S. GELLER and B. T. MATTHIAS: Superconductivity in the Ni–Pd–As system 156

Contents lists of Volume 2, Numbers 3 and 4 158

NUMBER 3

HELMUT BROSS und ALFRED SEEGER: Elektronentheoretische Untersuchungen über Fehlstellen in Metallen—V 161

A. C. DAMASK: Comparison of electron with neutron irradiation of α-brass 177

L. A. GIRIFALCO and J. R. STREETMAN: A theoretical calculation of the relaxation of atoms surrounding a vacancy in the body-centered cubic lattice 182

ROBERT F. BREBRICK: Deviations from stoichiometry in binary ionic crystals 190

ALLAN S. MILLER and ROBERT J. MAURER: Self-diffusion and electrical conductivity in silver bromide 196

HERBERT B. ROSENSTOCK: Infrared absorption by sodium chloride-type lattices 201
FREDERICK C. BROWN: Transient photoconductivity in AgCl at room temperature . . . 206
A. H. MONES and E. BANKS: Cation substitutions in $BaFe_{12}O_{19}$ 217
P. G. DE GENNES: Inelastic magnetic scattering of neutrons at high temperatures . . . 223
A. CHARLESBY and L. CALLAGHAN: Crystal distribution in various polyethylenes . . . 227
Book Review:
"Semiconductor Surface Physics" edited by R. H. KINGSTON (reviewed by C. HILSUM) . . 239
Books Received 240
Erratum 240

NUMBER 4

I. DZYALOSHINSKY: A thermodynamic theory of "weak" ferromagnetism of antiferromagnetics . 241
HERBERT B. CALLEN: A ferromagnetic dynamical equation 256
A. M. CLOGSTON: Susceptibility of the three-level maser 271
H. SUHL: Quantum analog of the ferromagnetic microwave amplifier 278
WOLFGANG LUDWIG: Über den Einfluss der Anharmonizität auf die Eigenschaften der Kristalle . 283
A. CHARLESBY and L. CALLAGHAN: Crystallinity changes in irradiated polyethylenes . . . 306
H. J. HROSTOWSKI and R. H. KAISER: Absorption spectrum of bismuth-doped silicon . . 315
Letter to the Editors:
B. T. MATTHIAS and S. GELLER: Superconductivity in the Pd–Se system 318
Book Review:
"Les Dislocations" by J. FRIEDEL (reviewed by W. M. LOMER) 320

AUTHOR INDEX TO VOLUME 4

(B) denotes a book review

ARGYRES, P. N. 19
BANKS, E. 217
BARSCH, GERHARD 27
BENSON, K. E. 154
BREBRICK, ROBERT F. 190
BROSS, HELMUT 161
BROWN, FREDERICK C. 206
CALLAGHAN, L. 227, 306
CALLEN, HERBERT B. 256
CHARLESBY, A. 227, 306
CLOGSTON, A. M. 271
DAMASK, A. C. 177
DANIEL, E. 111
DE GENNES, P. G. 71, 223
DOYLE, W. P. 144
DZYALOSHINSKY, I. 241
FELDTKELLER, ERNST 50
FRIEDEL, J. 71, 111
FULLER, C. S. 58, 155
GELLER, S. 154, 156, 318
GIRIFALCO, L. A. 182
HILSUM, C. 239 (B)
HROSTOWSKI, H. J. 148, 155, 315
JURETSCHKE, H. J. 118
KAISER, R. H. 148, 315
KOSEVICH, A. M. 1
KRUSEMEYER, H. J. 78
LAW, J. T. 91
LIFSHITZ, E. M. 1, 11
LOMER, W. M. 320 (B)
LUDWIG, WOLFGANG 283
MAITA, J. P. 68
MATTHIAS, B. T. 156, 318
MAURER, ROBERT J. 196
MILLER, ALLAN S. 196
MONES, A. H. 217
REISS, HOWARD 58
ROSENSTOCK, HERBERT B. 201
SEEGER, ALFRED 161
SMITH, CHARLES S. 128
STEINITZ, R. 118
STREETMAN, J. R. 182
SUHL, H. 278
TAUER, K. J. 135
THOMAS, D. G. 78
WALLIS, RICHARD F. 101
WEISS, R. J. 135
WERNICK, J. H. 154
WINDER, D. R. 128

VOL.
4
958

VOL.
4
195

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 1–10.

THEORY OF THE SHUBNIKOV—de HAAS EFFECT

E. M. LIFSHITS and A. M. KOSEVICH

Physical-Technical Institute, Academy of Science of the Ukranian S.S.R., Kark'off, U.S.S.R.

(*Received* 26 *December* 1956)

Abstract—Quantum oscillations of the electrical conductivity $\sigma^{\alpha\beta}$ and the resistivity $\rho^{\alpha\beta}$ tensors are studied using the general formulae of E. M. LIFSHITS.[1] It is shown that the oscillations of $\sigma^{\alpha\beta}$ and $\rho^{\alpha\beta}$ are expressed in terms of oscillations of the magnetic moment in the de Hass-van Alphen effect and the classical values of the mobility tensor. Asymptotes of the oscillation amplitudes in strong magnetic fields are investigated and the simplest cases in which complete calculations of the oscillation amplitudes may be carried out are discussed.

INTRODUCTION

A CONSISTENT quantum theory of the electrical conductivity of metals in magnetic fields was developed by E. M. LIFSHITS.[1] The relationship between the quantum kinetic equation and its classical analogue established by E. M. LIFSHITS[1] made it possible to determine quantum corrections to the classical values of the conductivity and in particular to find corrections which contain quantum oscillations of the conductivity.

This paper deals in greater detail with such oscillations (Shubnikov–de Haas effect[2a, b]).

As shown by E. M. LIFSHITS[1] the continuous (classical) part of the conductivity tensor may be written as

$$\left.\begin{aligned} \sigma^{\alpha\beta}_{\text{class}} &= -2\int\int \frac{\partial f_0}{\partial \epsilon}\chi^{\alpha\beta}m^*\,d\epsilon dp_z \qquad (\alpha,\beta \neq z,z) \\ \sigma^{zz}_{\text{class}} &= -2\int\int \frac{\partial f_0}{\partial \epsilon}(\chi^{zz}+\chi_0{}^{zz})m^*\,d\epsilon dp_z \end{aligned}\right\} \quad \text{(A)}$$

where $\chi(\epsilon, p_z)$ is found in a certain way from the value of the Green's function for the classical kinetic equation (see equation 5.9 of E. M. LIFSHITS[1]) and $f_0(\epsilon)$ is the Fermi distribution function.

The quantum oscillations of the conductivity tensor are expressed in terms of the quantities $\chi(\epsilon, p_z)$ by the following formulae:[1]

$$\Delta\sigma^{\alpha\beta} = \sum_{k=1}^{\infty} I_k{}^{\alpha\beta} \qquad (\alpha,\beta \neq z,z)$$

$$\Delta\sigma^{zz} = \sum_{k=1}^{\infty}(I_k{}^{zz} - L_k{}^{zz})$$

$$\left.\begin{aligned} I_k{}^{\alpha\beta} &= 2\int\int f_0 \frac{\partial(\chi^{\alpha\beta}m^*)}{\partial\epsilon}\hbar w^* \times \\ &\quad \exp(2\pi ikn)\,dndp_z \\ L_k{}^{zz} &= 2\int\int \frac{\partial f_0}{\partial\epsilon}\chi_0{}^{zz}m^*\hbar w \times \\ &\quad \exp(2\pi ikn)\,dndp_z \end{aligned}\right\} \quad \text{(B)}$$

where $\omega^* = eH/m^*c$, $m^* = (1/2\pi)(\partial S/\partial\epsilon)$ and $S = S(\epsilon, p_z)$ is an area of the section of the isoenergetic surface for an electron with an arbitrary dispersion law $\epsilon = \epsilon(\boldsymbol{p})$ by a plane $p_z = \text{const.}$ which is perpendicular to the magnetic field.

It is quite easy to separate out from the integrals in (B) the oscillating multipliers in which we are interested and to express the oscillations $\Delta\sigma^{\alpha\beta}$ in terms of oscillations of the magnetic moment and of the classical value of the mobility tensor.

1. *Oscillations of the conductivity*

We shall extract the oscillating part from the expressions for I_k and L_k in (B). We shall use the treatment of LIFSHITS and KOSEVICH[3] changing somewhat and simplifying the method used there.

We shall consider first the participation in the conductivity oscillations of only one group of electrons with a given dispersion law.

Instead of n, p_z we shall use variables n, ϵ in the integrals I_k and L_k changing correspondingly the order and the limits of integration and using the relationship

$$\hbar w^* \frac{\mathcal{J}(n, p_z)}{\mathcal{J}(n, \epsilon)} = 1 \Big/ \left(\frac{\partial n}{\partial p_z}\right)_\epsilon$$

which follows from the quasi-classical expression for the electron energy levels in a magnetic field:[3]

$$S(\epsilon, p_z) = (n+\gamma)\frac{ehH}{c}; \quad (0<\gamma<1), \quad n \gg 1.$$

For I_k we obtain

$$I_k = 2\int f_0(\epsilon) \sum \left\{ \int_{n_{\min}}^{n_{\max}} \frac{\partial(\chi m^*)}{\partial \epsilon} \left|\frac{\partial n}{\partial p_z}\right|_\epsilon^{-1} \exp(2\pi i k n)\, dn \right\} d\epsilon \qquad (1)$$

where Σ means summation over the intervals of the monotonic variation of $n(\epsilon, p_z)$ at a fixed value of ϵ.

Further calculations are based on the fact that, with the exception of $f_0(\varepsilon)$ and $\exp(2\pi i k n)$, all the quantities in the integrals I_k and L_k change very slowly as function of ε and n (compared with the intervals of variation of ε and n). Therefore for $n \gg 1$ the main contribution to the internal integral in (1) comes from integration in the region of the extremal points $n_m(\varepsilon)$. Near such points

$$n - n_m = \tfrac{1}{2}(p_z - p_z^m)^2 \left(\frac{\partial^2 n}{\partial p_z^2}\right)_m$$

where $n = n(\epsilon, p_z)$ and $n_m = n(\epsilon, p_z^m)$

and therefore

$$\left|\frac{\partial n}{\partial p_z}\right| = \sqrt{[2(\partial^2 n/\partial p_z^2)_m (n - n_m)]}.$$

The latter relationship makes it possible to obtain very simply the contribution to the internal integral in (1) by integration in the region of an extremal point:

$$\int^{n_m} \frac{\partial(\chi m^*)}{\partial \epsilon} \left|\frac{\partial n}{\partial p_z}\right|^{-1} \exp(2\pi i k n)\, dn$$

$$= \left|2\frac{\partial^2 n}{\partial p_z^2}\right|_m^{-\frac{1}{2}} \int^{n_m} \frac{\partial(\chi m^*)}{\partial \epsilon} \frac{\exp(2\pi i k n)}{\sqrt{|n - n_m|}}\, dn$$

$$\approx w_m(\epsilon) \exp(2\pi i k n_m) \int_{\pm\infty}^{0} \exp(2\pi i k x) \frac{dx}{\sqrt{|x|}}$$

$$= w_m(\epsilon) \exp(2\pi i k n_m) \frac{\exp(\pm i\pi/4)}{\sqrt{(2k)}}$$

where $$w_m(\epsilon) = \frac{\partial(\chi m^*)}{\partial \epsilon} \left|2\frac{\partial^2 n}{\partial p_z^2}\right|_m^{-\frac{1}{2}}.$$

The plus sign in the expression $\exp(\pm i\pi/4)$ corresponds to $n_{\min}$ and the minus sign to $n_{\max}$. At $n_{\min} = 0$ the corresponding term contributes nothing to the oscillating part of I_k and should be neglected. All the other extremal points are repeated twice (in the neighbouring intervals of the monotonic variation of $n(\varepsilon, p_z)$ at fixed ε), and we can therefore rewrite (1) as

$$I_k \approx 2\sqrt{\left(\frac{2}{k}\right)} \sum_m \exp(\pm i\pi/4) \times \times \int f_0(\epsilon) w_m(\epsilon) \exp[2\pi i k n_m(\epsilon)]\, d\epsilon \qquad (2)$$

where the summation applies to all the extremal points. In further discussion we shall omit the sign of the summation of the extremal points.

Using the form of $f_0(\varepsilon)$ near $\varepsilon = \xi$ and assuming $(dn_m/d\varepsilon)\,\xi \neq 0$, we find from (2)

$$I_k \approx 2\sqrt{\left(\frac{2}{k}\right)} w_m(\epsilon) \exp[2\pi i k n_m(\xi) \pm i\pi/4] \times \times \int_0^\infty f_0(\epsilon) \exp[2\pi i k (dn_m/d\xi)(\epsilon - \xi)]\, d\epsilon = \frac{\sqrt{2}}{\pi k^{3/2}} \frac{w_m(\epsilon)}{(dn_m/d\xi)} \frac{k\lambda}{\sinh k\lambda} \times \exp[2\pi i k n_m(\xi) - i\pi/2 \pm i\pi/4] \qquad (3)$$

where $\lambda = 2\pi^2\theta/\hbar\omega^*$ and ζ = chemical potential.

The oscillating part of $\sigma^{\alpha\beta}$ ($\alpha, \beta \neq z, z$) is given by the expression

$$\Delta\sigma^{\alpha\beta} = \sum_{k=1}^{\infty} \frac{\sqrt{2}}{\pi k^{3/2}} \cdot \frac{w_m{}^{\alpha\beta}(\xi)}{(dn_m/d\xi)} \times \left(\frac{k\lambda}{\sinh \cdot k\lambda}\right) \exp[2\pi i k n_m(\xi) - i(\pi/2 \pm \pi/4)]. \quad (4)$$

Expression (4) for $\Delta\sigma^{\alpha\beta}$ may be expressed in a more compact form if we introduce a quantity

$$F_k = 2\int\int f_0 \exp(2\pi i k n) 2\pi m^* \, d\epsilon dp_z. \quad (5)$$

It can be easily verified by a calculation similar to that given above that integration in the vicinity of the extremal points (excluding the trivial case $n_{\min} = 0$) results in the following expression for F_k

$$F_k = A_k(\xi) \exp[2\pi i k n_m(\xi) - i(\pi/2 \pm \pi/4)], \quad (6)$$

where

$$A_k(\xi) = \frac{2m^*(\xi, p_z{}^m) \cdot (k\lambda/\sinh k\lambda)}{k^{3/2}(dn_m/d\xi)|\partial^2 n/\partial p_z{}^2|_m{}^{\frac{1}{2}}} = 2\left(\frac{ehH}{kc}\right)^{3/2} \frac{m^*(\xi, p_z{}^m)(k\lambda/\sinh k\lambda)}{(dS_m/d\xi)|\partial^2 S(\xi, p_z)/\partial p_z{}^2|_m{}^{\frac{1}{2}}}.$$

In writing the above equality we used the relationship between the quantum number $n(\varepsilon, p_z)$. and the surface $S(\varepsilon, p_z)$. It follows that all the quantities which enter (4) or (6) are taken at the points $\varepsilon = \xi$, $p_z = p_z{}^m$ which correspond to the extremal values of the section of the bounding Fermi surface (these questions, as well as calculation of the integrals of the type I_k and F_k are dealt with in greater detail in.[3])

From the comparison of (3) and (4) it is clear that

$$I_k{}^{\alpha\beta} = \frac{1}{2\pi m^*} \frac{\partial(\chi^{\alpha\beta} m^*)}{\partial\epsilon}\bigg|_{\xi, p_z{}^m} \cdot F_k \quad (7)$$

Equation (4) may also be written as

$$\Delta\sigma^{\alpha\beta} = \frac{1}{2\pi m^*} \frac{\partial(\chi^{\alpha\beta} m^*)}{\partial\epsilon}\bigg|_{\xi, p_z{}^m} \cdot F; \quad F = \Sigma F_k. \quad (8)$$

We note that the oscillating part of the magnetic moment of the electron gas in the de Haas–van Alphen effect may also be written in terms of F

$$\Delta M^z = -h^{-3} \sum_k \int\int f_0 \frac{\partial\epsilon}{\partial H} \exp(2\pi i k n) \times 2\pi m^* \, d\epsilon dp_z \approx -S_m(\xi) F / h^3 H \frac{dS_m}{d\xi}$$

It follows that

$$F = -h^3 H \frac{d \ln S_m}{d\xi} \Delta M^z. \quad (9)$$

Substituting (9) in (8) we can express the oscillating quantity $\Delta\sigma^{\alpha\beta}$ in terms of the oscillating part of the magnetic moment ΔM^z.

$$\Delta\sigma^{\alpha\beta} = -\frac{h^3 H}{(dS_m/d\xi)} \times \frac{\partial}{\partial\xi}(\chi^{\alpha\beta} m^*) \frac{d \ln S_m}{d\xi} \Delta M^z. \quad (10)$$

We shall introduce a "classical mobility tensor" $q^{\alpha\beta}$ which is related to the classical conductivity tensor by

$$\sigma^{\alpha\beta}_{\text{class}} = 2h^{-3} \int\int f_0 q^{\alpha\beta} 2\pi m^* \, d\epsilon dp_z = N_0 \overline{q^{\alpha\beta}} \quad (11)$$

where N_0 = the number of electrons in the conduction band, $\bar{q}$ = mean value of the mobility taken with a statistical weight of f_0.

Since the classical expression (A) for $\sigma^{\alpha\beta}$ may be written as

$$\sigma^{\alpha\beta}_{\text{class}} = 2\int f_0 \cdot \frac{\partial(\chi^{\alpha\beta} m^*)}{\partial\epsilon} \, d\epsilon dp_z \quad (\alpha, \beta \neq z, z)$$

therefore for $q^{\alpha\beta}$ we find that

$$q^{\alpha\beta} = \frac{h^3}{2\pi m^*} \frac{\partial(\chi^{\alpha\beta} m^*)}{\partial\epsilon}. \quad (\alpha, \beta \neq z, z). \quad (12)$$

Substituting (12) in (10) we obtain the following expression for the oscillating part $\Delta\sigma^{\alpha\beta}$ (at $\alpha, \beta \neq z, z$) in terms of independent variables H, ξ:

$$(\Delta\sigma^{\alpha\beta})_{H,\xi} = -q_m{}^{\alpha\beta} H \frac{d \ln S_m}{d\xi} \Delta M^z. \quad (13)$$

If there are several groups of electrons which determine the conductivity of a metal, i.e. if there are several unoccupied overlapping energy bands, then each of the electron groups contributes to $\Delta\sigma^{\alpha\beta}$, so that $\Delta\sigma^{\alpha\beta}$ becomes a sum of components as in equation (13).

In a similar way we can obtain an expression for the oscillating part $\Delta\sigma^{\alpha\beta}$. Integrating $L_k{}^{zz}$ by the method used above

$$L_k{}^{zz} = \frac{2}{\sqrt{k}} \left(\frac{\chi_0{}^{zz} m^*}{\sqrt{(|\partial^2 n/\partial p_z{}^2|)}}\right)_{\varepsilon, p_z{}^m} \left(\frac{k\lambda}{\sinh k\lambda}\right) \times \exp[2\pi i k n_m(\xi) + i(\pi \pm \pi/4)]$$

we shall express the result in terms of F_k

$$L_k{}^{zz} = -ik\chi_{0m}{}^{zz}(\xi)\frac{dn_m}{d\xi}F_k. \tag{14}$$

Because of an additional multiplier $dn_m/d\xi$ the integral $L_k{}^{zz}$ (14) is much larger than $I_k{}^{zz}$ (in comparison with ξ/μ^*H, $\mu^*H \equiv \hbar\omega$)*. For this reason the main contribution to $\Delta\sigma^{zz}$ of the extremal sections for which $\chi_0{}^{zz}$ is not equal to zero is determined by

$$\Delta\sigma^{zz} = i\chi_{0m}{}^{zz}(\xi)\frac{dn_m}{d\xi}\sum_k kF_k.$$

Using the relationship of F with ΔM^z we can easily find for this case, in the same approximation, that

$$(\Delta\sigma^{zz})_{H,\xi} = \frac{h^3}{2\pi}H^2\chi_0{}^{zz}(\xi)\times \left(\frac{d\ln S_m}{d\xi}\right)^2\frac{\partial(\Delta M^z)}{\partial H}. \tag{15}$$

When $\chi_0{}^{zz}$ is zero at an extremal section (it is clear from the general formula, 59, E. M. LIFSHITS of[1] for $\chi_0{}^{zz}$ and from considerations of symmetry that this will take place, in particular, at the central section of the Fermi surface), it is essential to take into account in the integrals

$$Q_k(\epsilon) = \int^{n_m}\chi_0{}^{zz}m^*\left|\frac{\partial n}{\partial p_z}\right|^{-1}\exp(2\pi ikn)\,dn$$

the next (non-zero) terms of $\chi_0{}^{zz}$ written down in powers of $p_z - p_z{}^m$.

Assume that near an extremal point

$$\chi_0{}^{zz} = (p_z - p_z{}^m)^s\frac{1}{S!}\frac{\partial^s}{\partial p_z{}^s}\chi_0{}^{zz} \tag{16}$$

Then

$$Q_k(\epsilon) \approx (\pm 1)^s\frac{2^{(s-1)/2}}{S!}\frac{\partial^s\chi_0{}^{zz}}{\partial p_z{}^s}m^*\times \times\left|\frac{\partial^2 n}{\partial p_z{}^2}\right|_m^{-(s+1)/2}\cdot\int^{n_m}\left|n-n_m\right|^{(s-1)/2}\exp(2\pi ikn)\,dn \tag{17}$$

where, as before, the plus sign corresponds to the $n_{\min}$ and the minus sign to $n_{\max}$. To find the integral in (17) we may use the asymptotic formula for the integrals of this type which is given by LIFSHITS and KOSEVICH.[3]

If $\chi_0(\varepsilon, p_z)$ is a continuous function of p_z, then from $\chi_0(\varepsilon, p_z) = \chi_0(\varepsilon, -p_z)$ it follows that

$$\left.\frac{\partial\chi_0}{\partial p_z}\right|_{p_z=0} = 0,$$

and therefore for the central section the equation (16) begins with a quadratic term. This means that for the central section

$$Q_k(\epsilon) = \frac{m^*}{8\pi k^{3/2}}\frac{\partial^2\chi_0{}^{zz}}{\partial p_z{}^2}\left|\frac{\partial^2 n}{\partial p_z{}^2}\right|_0^{-3/2}\exp(2\pi ikn_m \pm i3/4\pi)$$

and the corresponding contribution to the oscillating part $\Delta\sigma^{zz}$ is given by

$$\Delta\sigma^{zz} = \frac{1}{2\pi}\left\{\frac{1}{m^*}\frac{\partial(\chi^{zz}m^*)}{\partial\epsilon} \mp \mp\frac{1}{2}\frac{dS_m/d\xi}{|\partial^2 S/\partial p_z{}^2|}\frac{\partial^2\chi_0{}^{zz}}{\partial p_z{}^2}\right\}_{\epsilon,p_z=0}\cdot F. \tag{18}$$

Introducing the component q^{zz} of the "classical mobility tensor" from a formula similar to (11), i.e.

$$q^{zz} = \frac{h^3}{2\pi m^*}\frac{\partial}{\partial\epsilon}(\chi^{zz}+\chi_0{}^{zz})m^* - \frac{h^3}{m^*}\frac{\partial}{\partial p_z}\frac{(\chi_0{}^{zz}m^{*2})}{\partial S/\partial p_z} \tag{19}$$

it can be easily seen that for central cross-section

$$q^{zz} = \frac{h^3}{2\pi m^*}\frac{\partial(\chi^{zz}m^*)}{\partial\epsilon} \mp h^3m^*\frac{\partial^2\chi_0{}^{zz}/\partial p_z{}^2}{2|\partial^2 S/\partial p_z{}^2|}.$$

and thus can write (18) as

$$(\Delta\sigma^{zz})_{H,\xi} = -q_m{}^{zz}H\left(\frac{d\ln S_m}{d\xi}\right)\Delta M^z. \tag{20}$$

If the bounding Fermi surface is convex, then the only extremal cross-section is the central one, and therefore the oscillating parts of the electrical conductivity of a given electron group are given by (13) and (20), and all the components of $\Delta\sigma^{\alpha\beta}$ are of the same order in oscillation amplitudes. If in addition to the central one, there are other extremal cross-sections, then $\Delta\sigma^{zz}$ is given by (15) and the amplitude of the oscillations of $\Delta\sigma^{zz}$ will be considerably greater than the amplitude of the oscillations of $\Delta\sigma^{\alpha\beta}(\alpha, \beta \neq z, z)$ for the same electron group.

The final formulae (13), (15) and (20) for the oscillating part $\Delta\sigma^{\alpha\beta}$ are given in terms of independent variables ξ and H. As shown by E. LIFSHITS,[5] in concrete applications of these formulae and on comparison with experimental results we must take into account the oscillations of the chemical potential $\xi = \xi(H)$ which follow from the constancy of the number of electrons in all bands:

VOL 4 195

$$2h^{-3}\sum_n\sum\int f_0 2\pi m_j^*\,dp_z\Delta\epsilon_n$$

$$\equiv\sum_j N_j = N = \text{const.}$$

(summation in j applies to all bands).

In contrast to the de Haas–van Alphen effect when these oscillations may be neglected,[3] in our case, due to the large value of $\sigma^{\alpha\beta}$, they are important.

Denoting by $N_0(\xi)$ the classical dependence of the number of electrons on the chemical potential

$$N^0(\xi) = 2h^{-3}\sum_j\int\int f_0 2\pi m_i^*\,dp_z d\epsilon$$

$$= 2h^{-3}\sum_j\int f_0(d\mathbf{p})$$

we find, using the Poisson formula and (5) above,

$$N(\xi) = N^0(\xi)+h^{-3}\sum_j\sum_{k=-\infty}^{\infty}F_k{}^j$$

$$= N^0(\xi)+h^{-3}\sum_j F^j(\xi). \qquad (21)$$

Taking $\xi = \xi_0+\Delta\xi$, where ξ_0 = chemical potential at $H = 0$, we write

$$N^0(\xi) = N^0(\xi_0)+\sum_j\frac{\partial N_j{}^0}{\partial\xi_0}\Delta\xi$$

and it then follows from (21) that

$$\Delta\xi\cdot h^3\sum_j\frac{\partial N_j{}^0}{\partial\xi} = -\sum_j F^j(\xi_0). \qquad (22)$$

Assuming that the oscillating correction to $\sigma^{\alpha\beta}$ due to the dependence $\xi = \xi(H)$ has the form

$$\Delta\sigma_1{}^{\alpha\beta} = \frac{\partial\sigma^{\alpha\beta}}{\partial\xi_0}\Delta\xi \qquad (23)$$

we substitute into (23) the value of $\Delta\xi$ from (22):

$$\Delta\sigma_1{}^{\alpha\beta} = -\left(\frac{\partial\sigma^{\alpha\beta}}{\partial\xi_0}\sum_j F^j(\xi_0)\right)\Big/h^3\sum_j\frac{\partial N_j{}^0}{\partial\xi_0}.$$

If we now use (9) then the oscillations $\Delta\sigma_1{}^{\alpha\beta}$ in terms of ΔM^z are given by:

$$\Delta\sigma_1{}^{\alpha\beta} = \left(H\frac{\partial\sigma^{\alpha\beta}}{\partial\xi_0}\sum_j\frac{d\ln S_{mj}}{\partial\xi_0}\Delta M_j{}^z\right)\Big/\sum_k\frac{\partial N_k{}^0}{\partial\xi_0} \qquad (24)$$

Finally, $\sigma^{\alpha\beta}$ in (24) may be expressed in terms of the classical "mobility tensor" (11):

$$\Delta\sigma_1{}^{\alpha\beta} = (H\sum_k\frac{\partial(N_k{}^0\overline{q_k{}^{\alpha\beta}})}{\partial\xi}\sum_j\frac{d\ln S_{mj}}{\partial\xi_0}\Delta M_j{}^z)/\sum_k\frac{\partial N_k{}^0}{\partial\xi_0}. \qquad (25)$$

The experimentally observed conductivity oscillations on variation of the magnetic field are given by the sum $\Delta\sigma^{\alpha\beta}+\Delta\sigma_1{}^{\alpha\beta}$, with $\xi = \xi_0$ in the formulae (13), (15) and (20) (or those which correspond to them).

For $\alpha, \beta \neq z, z$ this sum has the form

$$\Delta\sigma^{\alpha\beta}+\Delta\sigma_1{}^{\alpha\beta} = H\sum_j\left[\left(\sum_k\frac{\partial(N_k{}^0\overline{q_k{}^{\alpha\beta}})}{\partial\xi_0}\right)\Big/\sum_k\frac{\partial N_k{}^0}{\partial\xi_0}-q_{mj}{}^{\alpha\beta}(\xi_0)\right]\frac{d\ln S_{mj}}{\partial\xi_0}\Delta M_j{}^z. \qquad (26)$$

If all the Fermi surfaces are convex then $\Delta\sigma^{zz}+\Delta\sigma_1{}^{zz}$ are given by a similar expression

$$\Delta\sigma^{zz}+\Delta\sigma_1{}^{zz} = H\sum_j\left[\left(\sum_k\frac{\partial(N_k{}^0\overline{q_k{}^{zz}})}{\partial\xi_0}\right)\Big/\sum_k\frac{\partial N_k{}^0}{\partial\xi_0}-q_{mj}{}^{zz}(\xi_0)\right]\frac{d\ln S_{mj}}{d\xi_0}\Delta M_j{}^z. \qquad (27)$$

When the Fermi surface has non-central sections, then

$$\Delta\sigma^{zz}+\Delta\sigma_1{}^{zz} = (h^3/2\pi)H^2\sum_i\chi_{0mi}{}^{zz}(\xi_0)(d\ln S_{mi}/d\xi_0)^2\frac{\partial(\Delta M_i{}^z)}{\partial H} \qquad (28)$$

where summation applies to all the Fermi surfaces with non-central sections.

From the formulae for the oscillating part of the conductivity $\Delta\sigma^{\alpha\beta}$ (we shall now use $\Delta\sigma^{\alpha\beta}$ to denote all terms of $\sigma^{\alpha\beta}$ which oscillate with variation of the magnetic field, i.e. $\Delta\sigma^{\alpha\beta}$ shall denote sums of the form given in equations, 26 and 27) it is clear that each electron group contributes to $\Delta\sigma^{\alpha\beta}$. It is found that the contribution of each group is related only to ΔM^z of the corresponding electron group. Therefore the oscillation period of $\Delta\sigma^{\alpha\beta}$ is always determined by the same multiplier as the period of the oscillations of ΔM^z and, in the de Haas-van Alphen effect is identical with the latter[3] given by:

$$\Delta\left(\frac{1}{H}\right)=\frac{eh}{cS_m(\xi_0)}.$$

On the subject of the amplitude of the oscillations $\Delta\sigma^{\alpha\beta}$ we should note the following:

Firstly, in contrast to the oscillations of ΔM^z the amplitude of that component of $\Delta\sigma^{\alpha\beta}$ which corresponds to a given "anomalously small" band[3] is determined by all, including the "normal", electron groups.

Secondly, the same unknown quantities $\chi(\epsilon, p_z)$ which determine the classical conductivity enter the expressions for the amplitude.

In certain special cases given below in Section 4 the quantities $\chi(\epsilon, p_z)$ may be found by solving the appropriate kinetic equation and the oscillation amplitudes may be calculated completely.

2. *Asymptotic form of the conductivity oscillations in strong magnetic fields*

We shall elucidate the behaviour of the oscillation amplitudes of $\Delta\sigma^{\alpha\beta}$ in strong magnetic fields when $\gamma_0 \ll 1$, where, as stated by E. LIFSHITS, *et al.*[4] $\gamma_0 = 1/(eH/m_0ct_0)$ and mass m_0 time t_0 are certain characteristic quantities. In this case we may write down the amplitudes in an asymptotic form in powers of γ_0, using asymptotes for $\sigma_0{}^{\alpha\beta}=\sigma^{\alpha\beta}_{\text{class}}$ found by E. M. LIFSHITS[1] and LIFSHITS, AZBEL', and KAGANOV.[4]

If the bounding Fermi surface consists of several *closed* surfaces, then the asymptotes of $\sigma_0{}^{\alpha\beta}$ are the form

$$\sigma_0{}^{\alpha\beta}(\boldsymbol{H}) = \begin{bmatrix} \gamma_0{}^2a_{11} & \gamma_0a_{12}+\gamma_0{}^2a'_{12} & \gamma_0a_{13} \\ \gamma_0a_{21}+\gamma_0{}^2a'_{21} & \gamma_0{}^2a_{22} & \gamma_0a_{23} \\ \gamma_0a_{31} & \gamma_0a_{32} & a_{33} \end{bmatrix}. \quad (29)$$

Here $a_{\alpha\beta}(\xi_0)$ is a matrix, decomposition of whose elements in powers of γ_0 begins with zero terms, and

$$\gamma_0a_{12}=\frac{ec}{H}\Big(\sum_i N_i{}^+-\sum_k N_k{}^-\Big) \quad (30)$$

where N^+ = "electron" density, N^- = "hole" density for the corresponding groups.

Since for each electron group $\sigma_0{}^{\alpha\beta}=N_0\overline{q^{\alpha\beta}}$, then asymptotic decomposition of the elements of $\overline{q^{\alpha\beta}}$ begins with terms of the same order relative to γ_0 as in $\sigma^{\alpha\beta}$, that is

$$\overline{q^{\alpha\beta}} = \begin{bmatrix} \gamma_0{}^2\tilde{u}_{11} & \dfrac{ec}{H}+\gamma_0{}^2\tilde{u}_{12} & \gamma_0\tilde{u}_{13} \\ -\dfrac{ec}{H}+\gamma_0{}^2\tilde{u}_{21} & \gamma_0{}^2\tilde{u}_{22} & \gamma_0\tilde{u}_{23} \\ \gamma_0\tilde{u}_{31} & \gamma_0\tilde{u}_{32} & \tilde{u}_{33} \end{bmatrix}. \quad (31)$$

Decomposition of the elements of the matrix $\tilde{u}_{\alpha\beta}$ in powers of γ_0 begins, in general, with the zero term (for special cases of decomposition of $a_{\alpha\beta}$ and $\tilde{u}_{\alpha\beta}$ see LIFSHITS, AZBEL', and KAGANOV).[4] All elements of $\tilde{u}_{\alpha\beta}$ depend on the form of the collision integral and are, in general, functions of ξ_0.

Since $q^{\alpha\beta}{}_m$ is in general of the same order of magnitude as $\overline{q^{\alpha\beta}}$, then the asymptotes of $q_m{}^{\alpha\beta}$ in strong fields, should be of the form given in (31). The possibility of accurate calculation of the first term of the asymptotic decomposition of $q_m{}^{xy}$ is of interest. This possibility is considered below.

In crossed fields, with the magnetic field $\boldsymbol{H}$ along z-axis and the electric field $\boldsymbol{E}$ along the y-axis, an electron is in a steady state with a non-zero average constant velocity. This velocity is directed along the x-axis and is equal to cE/H.* This velocity depends neither on the dispersion law for electrons nor on the state of the electron, i.e. it does not change on collision. A steady-state electron mobility $q_0{}^{xy}=ec/H$ is related to this electron velocity.

* This can be easily verified if we remember that in the steady state the average value of the Lorentz force which acts on the electron is zero. Since the Lorentz force does not depend on the form of the dispersion law, therefore the steady-state electron velocity in crossed fields is the same for any dispersion law.

In strong magnetic fields, when the mean time interval between the electron collisions is much larger than the period of electron revolution in a classical orbit, q_0^{xy} is the main part of the element q^{xy} of the mobility tensor and therefore

$$q_m^{xy} = \frac{ec}{H} + \gamma_0^2 u_{m1}^2 \tag{32}$$

where $\gamma_0^2\ u_m^{12}$ = mobility specified by the collision integral. Comparing (32) and (31) we find that the first terms of q^{xy} and q_m^{xy} written down in terms of powers of γ_0, are identical.

From (26) and (27), and also from (31) and (32) it follows that the oscillation amplitudes of all elements of $\Delta\sigma^{\alpha\beta}$, except $\Delta\sigma^{xy}$, are of the same order in γ_0 as the corresponding elements of $\sigma_0^{\alpha\beta}$. The magnitude of $\Delta\sigma^{xy}$ is one order in γ_0 higher than σ_0^{xy} if $\sum_k N_k^+ \neq \sum_i N_i^+$. We can therefore write for the relative magnitude of the oscillations, for convex Fermi surfaces and

$$\sum_k N_k^+ \neq \sum_i N_i^-,$$

$$\Delta\sigma^{\alpha\beta}/\sigma_0^{\alpha\beta} = H \sum_j \Psi_j^{\alpha\beta} \frac{d \ln S_{mj}}{d\xi_0} \Delta M_j^z \tag{33}$$

where

$$\psi^{\alpha\beta} = \begin{bmatrix} \psi_{11} & \gamma_0\psi_{12} & \psi_{13} \\ \gamma_0\psi_{21} & \psi_{22} & \psi_{23} \\ \psi_{31} & \psi_{32} & \psi_{33} \end{bmatrix} \tag{34}$$

and the matrix $\Psi_{\alpha\beta}$ is

$$\psi_{\alpha\beta}^j = \left(\frac{\partial \ln a_{\alpha\beta}}{\partial \xi_0} \Big/ \sum_k \frac{\partial N_k^0}{\partial \xi_0} \right) - u_{mj}^{\alpha\beta} \Big/ a_{\alpha\beta}$$

$$(\alpha, \beta \neq x, y)$$

$$\psi_{xy}^j = \left(\frac{\partial a'_{12}}{\partial \xi_0} \Big/ a_{12} \sum_k \frac{\partial N_k^0}{\partial \xi_0} \right) - u_{mj}^{12}/a_{12}$$

and consequently the elements of the matrix $\psi_{\alpha\beta}$ when written down in powers of γ_0 begin with the zero term in γ_0.

It is interesting to note that the relative magnitude of the oscillations of σ^{xy} is one order (of γ_0) lower than the relative magnitude of the oscillations or remaining elements of the tensor $\sigma^{\alpha\beta}$.* This is due to a simple physical reason that the asymptotic value of σ_0^{xy}, in strong fields and when $\sum_k N_k^+ \neq \sum_i N_i^-$ in equation (30), is, for a given metal, a general quantity which does not undergo quantum oscillations.

In the case $\sum_k N_k^+ = \sum_i N_i^-$ the expression for σ_0^{xy} starts with γ_0^2 and therefore Ψ^{xy} begins with a zero term: $\Psi^{xy} = \Psi'_{12}$, where

$$\psi'_{12} = \left(\frac{\partial \ln a'_{12}}{\partial \xi_0} \Big/ \sum_k \frac{\partial N_k^0}{\partial \xi_0} \right) - u_m^{12}/a'_{12}.$$

The other terms of $\psi^{\alpha\beta}$ are, as before, given by (34).

3. *Resistance oscillations*

In experiments the electrical conductivity tensor $\sigma^{\alpha\beta}$ is not usually measured, but the resistivity tensor $\rho^{\alpha\beta} = \sigma_{\alpha\beta}^{-1}$ and therefore we must find the oscillating part of $\rho^{\alpha\beta}$.

The relationship between the elements of the tensors $\rho^{\alpha\beta}$ and $\sigma^{\alpha\beta}$ is given by

$$\rho^{\alpha\beta} = \mathcal{D}_{\beta\alpha}/||\sigma||; \quad ||\sigma|| = \det|\sigma^{\alpha\beta}| \tag{35}$$

where $\mathcal{D}_{\alpha\beta}$ = algebraic complement of the element $\sigma^{\alpha\beta}$ in the determinant $||\sigma||$.

We shall write $\sigma^{\alpha\beta} = \sigma_0^{\alpha\beta} + \Delta\sigma^{\alpha\beta}$, where $\Delta\sigma^{\alpha\beta}$, the oscillating part of $\sigma^{\alpha\beta}$, is a small quantum correction to $\sigma_0^{\alpha\beta}$. Leaving in only the terms linear in $\Delta\sigma^{\alpha\beta}$ we easily find that

$$\begin{aligned} ||\sigma|| &= ||\sigma_0|| + \mathcal{D}_{\alpha\beta}^0 \Delta\sigma^{\alpha\beta} \\ &= ||\sigma_0||(1 + \rho_0^{\alpha\beta}\Delta\sigma^{\alpha\beta}) \end{aligned} \tag{36}$$

(the sign of summation from 1 to 3 of the double indexes is omitted).

Similarly in the same approximation we find

$$\mathcal{D}_{\alpha\beta} = \mathcal{D}_{\alpha\beta}^0 + \epsilon_{\alpha kl}\epsilon_{\beta pq}\sigma_0^{kp}\Delta\sigma^{lq} \tag{37}$$

where ϵ_{ikl} = unit antisymmetric tensor of the third order.

If (36) and (37) are used, then from (35) it follows that

$$\Delta\rho^{\alpha\beta} = (||\rho_0||\epsilon_{\beta kl}\epsilon_{\alpha pq}\sigma_0^{kp} - \rho_0^{\alpha\beta}\rho_0^{lq})\Delta\sigma^{lq}$$

$$\Delta\rho^{\alpha\beta}(\boldsymbol{H}) = \Delta\rho^{\beta\alpha}(-\boldsymbol{H}),$$

$$\text{where} \quad ||\rho_0|| = \det|\rho_0^{\alpha\beta}| = ||\sigma_0||^{-1}. \tag{38}$$

From (38) we see that in its general form the formula for the oscillating part of the electrical resistivity tensor is very complex. The expression for $\Delta\rho^{\alpha\beta}$ contains both the classical values and the oscillating parts of, generally speaking, all components of $\sigma^{\alpha\beta}$. Even in very simple cases $\Delta\rho^{\alpha\beta}$

* In certain special cases, in particular for the case of the isotropic dispersion law, Ψ^{xy} may start with terms of the order γ_0^2 (at $\sum_k N_k^+ \neq \sum_i N_i^-$).

contains a number of components which may possess amplitudes of the same order and different periods of oscillation. Simplification of (38) can be only achieved in a few special cases.

In particular if there are Fermi surfaces with non-central sections (it is clear that there is always an even number of such sections, and those which are symmetrical with respect to the $p_z = 0$ section contribute the same amount to the oscillations of $\Delta\sigma^{\alpha\beta}$) at which $\chi_0{}^{zz} \neq 0$, then, at magnetic fields which satisfy $\gamma_0 \sim 1$, we can leave in (38) only those components which contain $\Delta\sigma^{zz}$. The components of $\Delta\rho^{\alpha\beta}$ become then

$$\Delta\rho^{xx} = (||\rho_0||\sigma_0{}^{yy} - \rho_0{}^{xx}\rho_0{}^{zz})\Delta\sigma^{zz};$$
$$\Delta\rho^{yy} = (||\rho_0||\sigma_0{}^{xx} - \rho^{yy}\rho_0{}^{zz})\Delta\sigma^{zz};$$
$$\Delta\rho^{xy} = (||\rho_0||\sigma_0{}^{xy} - \rho_0{}^{xy}\rho_0{}^{zz})\Delta\sigma^{zz};$$
$$\Delta\rho^{\alpha z} = -\rho_0{}^{\alpha z}\rho_0{}^{zz}\Delta\sigma^{zz} \quad (\alpha = x, y, z)$$

where $\Delta\sigma^{zz}$ is found from (28).

In the case of only convex Fermi surfaces we can find the asymptotes of $\Delta\rho^{\alpha\beta}$ in strong magnetic fields, ($\gamma_0 \ll 1$). For this purpose, as in Section 2, we shall use the asymptotes of the tensors $\sigma_0{}^{\alpha\beta}$ (29) and $\rho_0{}^{\alpha\beta(4)}$ which make it possible to represent the relative magnitudes of the oscillations of the resistivity tensor in the form

$$\frac{\Delta\rho^{\alpha\beta}}{\rho_0{}^{\alpha\beta}} = H\sum_j \Phi_j{}^{\alpha\beta}\frac{d\ln S_{mj}}{d\xi_0}\Delta M_j{}^z \qquad (39)$$

In the case $\sum_k N_k{}^+ \neq \sum N_i{}^-$ the matrix $\Phi^{\alpha\beta}$ is given by

$$\Phi^{\alpha\beta} = \begin{bmatrix} \phi_{11} & \gamma_0\phi_{12} & \phi_{13} \\ \gamma_0\phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix}. \qquad (40)$$

When the matrix $\phi_{\alpha\beta}$ is written in powers of γ_0, the first term is of the zero order.

It is found that, as in the case of the conductivity oscillations, the relative magnitude of the oscillations of the element ρ^{xy} is one order (of γ_0) lower than the relative magnitudes of the oscillations of the remaining elements of the tensor $\rho^{\alpha\beta}$.

When $\sum_k N_k{}^+ = \sum N_i{}^-$ elements of Φ^{xy} also begin with the zero order in and the remaining elements of the $\Phi^{\alpha\beta}$ matrix are of the previous form (40).

4. *Calculations of the oscillations in certain special cases*

In this section we shall discuss some very simple cases, when simplification of the general formulae for the oscillating parts of the tensors $\sigma^{\alpha\beta}$ and $\rho^{\alpha\beta}$ is possible.

The relationship for $\Delta\sigma^{\alpha\beta}$ is greatly simplified when only one conduction band is present and the Fermi surface is convex. We then have only one component in (26) and (27) and we obtain

$$\Delta\sigma^{\alpha\beta} =$$
$$= H\left\{\overline{q^{\alpha\beta}} - q_m{}^{\alpha\beta}(\xi_0) - \frac{\partial\overline{q^{\alpha\beta}}}{\partial\xi_0}\Big/\frac{\partial\ln N^0}{\partial\xi_0}\right\}\frac{d\ln S_m}{d\xi_0}\Delta M^z. \qquad (41)$$

If $q^{\alpha\beta}$ does not depend on ϵ and p_z, the sum (41) becomes zero. This means that the oscillations of $\sigma^{\alpha\beta}$ and $\rho^{\alpha\beta}$ in the case of the single-band conduction and the convex Fermi surface are completely determined by the dependence of the mobility on ϵ and p_z.

To calculate the amplitudes of the oscillations $\Delta\sigma^{\alpha\beta}$ and $\Delta\rho^{\alpha\beta}$, in the case of both one and several electron groups, it is necessary to know the functions $\chi(\epsilon, p_z)$ which may be obtained only by making some assumptions about the collision integral.

If the collision integral can be replaced in the kinetic equation by a "relaxation time" t_0 (which corresponds to $\tilde{W} = 1$ in the equation (8) stated by E. M. LIFSHITS),[1] then

$$\chi^{\alpha\beta} = 2\pi\frac{e^2t_0}{h^3}2\,\mathrm{Re}\sum_k \frac{\gamma}{\gamma + ik}v_{-k}{}^\alpha v_k{}^\beta;$$

$$\chi_0{}^{zz} = 2\pi\frac{e^2t_0}{h^3}(v_0{}^z)^2; \quad \gamma = \frac{1}{t_0w^*} = \frac{m^*c}{et_0H}, \qquad (42)$$

where $\mathbf{V}_k$ = Fourier component of the electron velocity.

In the isotropic case equation (42) is further simplified. Noting that for an arbitrary isotropic dispersion law $\epsilon = \epsilon(p)$ the relationship $\boldsymbol{p} = m^*\boldsymbol{v}$, $\boldsymbol{v} = \nabla_{\boldsymbol{p}}\epsilon$† holds, we find that

† In the isotropic case $\mathbf{v} = (\partial\varepsilon/\partial p)\,.\,(\mathbf{P}/p)$ and from $S(\varepsilon, p_z) = \pi(p^2(\varepsilon) - p_z{}^2)$ it follows that $m^* = p(\partial p/\partial\varepsilon) = p/\partial\varepsilon/\partial p$, therefore $\mathbf{P} = m^*\mathbf{v}$.

$$\chi^{xx}=\chi^{yy}=\frac{e^2t_0}{h^3}\frac{\gamma^2}{1+\gamma^2}\frac{S(\epsilon,p_z)}{m^{*2}};$$

$$\chi^{xy}=\frac{e^2t_0}{h^3}\frac{\gamma}{1+\gamma^2}\frac{S(\epsilon,p_z)}{m^{*2}};$$

$$\chi^{\alpha z}=0 \text{ where } \alpha=x,y,z; \tag{43}$$

$$\chi_0{}^{zz}=2\pi\frac{e^2t_0}{h^3}\frac{{p_z}^2}{m^{*2}};$$

$$S(\epsilon,p_z)=\pi(p^2(\epsilon)-{p_z}^2)$$

where $p=p(\epsilon)$ is the inverse of $\epsilon=\epsilon(p)$.

From (43) and determination of the average value of the "classical mobility tensor" (11) on the assumption of a δ-type dependence $f_0'(\epsilon)$ it follows that

$$\overline{q^{xx}}=\overline{q^{yy}}=\frac{q_0}{1+w^{*2}{t_0}^2};\quad \overline{q^{xy}}=\frac{w^*t_0q_0}{1+w^{*2}{t_0}^2}$$

$$\overline{q^{zz}}=q_0=\frac{e^2t_0}{m^*};\quad \overline{q^{xz}}=\overline{q^{yz}}=0;\quad t=t_0(\xi_0) \tag{44}$$

where as before $\omega^*=eH/m^*c$.

From (12) and (19), using (43) we find

$$q_m{}^{\alpha\beta}(\xi_0)=\overline{q^{\alpha\beta}}+\frac{p^2(\xi_0)}{2m^*}\frac{\partial q^{\alpha\beta}}{\partial\xi_0}\quad(\alpha,\beta\neq z,z)$$
$$q_m{}^{zz}=\overline{q^{zz}} \tag{45}$$

We shall use (44) and (45) to consider two examples.

(*a*) *Single conduction band.* When only one conduction band is present and $\partial \ln N^0/\partial\xi_0=3m^*/p^2(\xi_0)$ the oscillating parts $\Delta\sigma^{\alpha\beta}$ are of the form

$$\Delta\sigma^{\alpha\beta}=-\tfrac{1}{3}H\frac{\overline{\partial q^{\alpha\beta}}}{\partial\xi_0}\Delta M^z\quad(\alpha,\beta\neq z,z) \tag{46}$$
$$\Delta\sigma^{zz}=\tfrac{2}{3}H\frac{\overline{\partial q^{zz}}}{\partial\xi_0}\Delta M^z$$

where $\overline{q^{\alpha\beta}}$ is found from (44).

In strong magnetic fields ($\gamma\ll 1$) all the elements of $\overline{q^{\alpha\beta}}$ and their derivatives can be easily expressed in terms of q_0 and we obtain†

$$\frac{\Delta\sigma^{xx}}{\sigma_0{}^{xx}}=\frac{\Delta\sigma^{yy}}{\sigma_0{}^{yy}}=\tfrac{1}{3}H\frac{\partial\ln q_0}{\partial\xi_0}\frac{\Delta M^z}{N_0}$$
$$\frac{\Delta\sigma^{xy}}{\sigma_0{}^{xy}}=-\tfrac{2}{3}\gamma^2H\frac{\partial\ln q_0}{\partial\xi_0}\frac{\Delta M^z}{N^0}; \tag{47}$$
$$\frac{\Delta\sigma^{zz}}{\sigma_0{}^{zz}}=\tfrac{2}{3}H\frac{\partial\ln q_0}{\partial\xi_0}\frac{\Delta M^z}{N_0}.$$

The expression for the oscillating part of the magnetic moment ΔM^z in the case of an isotropic dispersion law is found from a general formula,[3] in which we shall substitute

$$S_m(\xi_0)=\pi p^2(\xi_0);$$
$$\frac{dS_m}{d\xi}=2\pi m^*;\qquad \frac{\partial^2 S}{\partial {p_z}^2}=2\pi.$$

Using $\sigma_0{}^{\alpha\beta}=N_0\overline{q^{\alpha\beta}}$ and the equations (38) and (46) we can calculate the oscillations of the resistivity tensor:

† If $m^*=m=$ const, then

$$\frac{\partial\ln q_0}{\partial\xi_0}-\frac{\partial\ln t_0}{\partial\xi_0},$$

and the expression for $\Delta\sigma^{xy}/\sigma_0{}^{xy}$ in (47) is identical with that stated by E. LIFSHITS[5] where, however, an error occurred in the numerical multiplier.

$$\Delta\rho^{xx}=\Delta\rho^{yy}=\frac{H}{3{\sigma_0}^2}\left\{(1-w^{*2}{t_0}^2)\frac{\partial\overline{q^{xx}}}{\partial\xi_0}-2w^*t_0\frac{\partial\overline{q^{xy}}}{\partial\xi_0}\right\}\Delta M^z$$

$$\Delta\rho^{xy}=\frac{H}{3{\sigma_0}^2}\left\{(1+3w^{*2}{t_0}^2)\frac{\partial\overline{q^{xy}}}{\partial\xi_0}-2w^*t_0\frac{\partial\overline{q^{xx}}}{\partial\xi_0}\right\}\Delta M^z \tag{48}$$

$$\Delta\rho^{zz}=-\frac{2H}{3\sigma_0{}^2}\left\{2w^*t_0\frac{\partial\overline{q^{xy}}}{\partial\xi_0}+\frac{\partial\overline{q^{zz}}}{\partial\xi_0}\right\}\Delta M^z$$

where $\sigma_0 = N_0(e^2 t_0/m^*)$ is the conductivity of a metal in the absence of a magnetic field.

In strong magnetic fields we can assume $\omega^* t_0 \gg 1$, which changes (48) into the following expression

$$\frac{\Delta\rho^{xx}}{\rho_0{}^{xx}} = \frac{\Delta\rho^{yy}}{\rho_0{}^{yy}} = \tfrac{1}{3}H\frac{\partial \ln q_0}{\partial \xi_0}\frac{\Delta M^z}{N_0}$$
$$\frac{\Delta\rho^{xy}}{\rho_0{}^{xy}} = -\tfrac{8}{3}\gamma^2 H\frac{\partial \ln q_0}{\partial \xi_0}\frac{\Delta M^z}{N_0}; \qquad (49)$$
$$\frac{\Delta\rho^{zz}}{\rho_0{}^{zz}} = -\tfrac{2}{3}H\frac{\partial \ln q_0}{\partial \xi_0}\frac{\Delta M^z}{N_0}.$$

The relative magnitudes of the oscillations $\Delta\sigma^{xy}/\sigma_0{}^{xy}$, and also $\Delta\rho^{xy}/\rho_0{}^{xy}$, in strong magnetic fields are in this case two orders of γ_0 lower than the relative magnitudes of the oscillations of the remaining elements of the tensors $\sigma^{\alpha\beta}$ and $\rho^{\alpha\beta}$.

(*b*) *Two bands with* $N^+ = N^-$. We shall assume that there are two electron groups with a quadratic dispersion law, and that for one of them ("electrons") $(1/2\pi)(dS_{m1}/d\epsilon) = m_1 > 0$ and for the other $(1/2\pi)(dS_{m2}/)d\epsilon) = m_2 < 0$ ("holes") and that $N^+ = N^-$.

Denoting the "electron" mobility by q_1 and the "hole" mobility by q_2, we easily find from the general equations (26) and (27) that

$$\Delta\sigma^{\alpha\beta} = -\frac{H}{(1+\mu)\xi_0}\left\{\left[\mu(\overline{q_1{}^{\alpha\beta}} - \overline{q_2{}^{\alpha\beta}}) + (\mu+\tfrac{1}{3})\xi_0\frac{\partial\overline{q_1{}^{\alpha\beta}}}{\partial\xi_0} + \tfrac{2}{3}\xi_0\frac{\partial\overline{q_2{}^{\alpha\beta}}}{\partial\xi_0}\right]\Delta M_1{}^z + \left[\mu(\overline{q_1{}^{\alpha\beta}} - \overline{q_2{}^{\alpha\beta}}) + \tfrac{2}{3}\mu\xi_0\frac{\partial\overline{q_1{}^{\alpha\beta}}}{\partial\xi_0} - (\tfrac{2}{3}\mu+1)\xi_0\frac{\partial\overline{q_2{}^{\alpha\beta}}}{\partial\xi_0}\right]\Delta M_2{}^z\right\} \quad (\alpha,\beta \neq z, z) \qquad (50)$$

$$\Delta\sigma^{zz} = -\frac{H}{(1+\mu)\xi_0}\left\{\left[\mu(\overline{q_1{}^{zz}} - \overline{q_2{}^{zz}}) - \tfrac{2}{3}\xi_0\left(\frac{\partial\overline{q_1{}^{zz}}}{\partial\xi_0} - \frac{\partial\overline{q_2{}^{zz}}}{\partial\xi_0}\right)\right]\Delta M_1{}^z + \left[\mu(\overline{q_1{}^{zz}} - \overline{q_2{}^{zz}}) + \tfrac{2}{3}\mu\xi_0\left(\frac{\partial\overline{q_1{}^{zz}}}{\partial\xi_0} - \frac{\partial\overline{q_2{}^{zz}}}{\partial\xi_0}\right)\right]\Delta M_2{}^z\right\} \qquad (51)$$

where $\mu = m_2/m_1$ and the mobilities are given by (44), where we substitute $m^* = m_1$ for the "electrons" and $m^* = -m_2$ for the "holes".

In the limiting cases $\mu \gg 1$ or $\mu \ll 1$ certain terms in (50) and (51) may be omitted, but the general character of the expressions is not altered.

The equations (50) and (51) show that the assumption $N^+ = N^-$ does not materially alter the general equations for $\Delta\sigma^{\alpha\beta}$.

Translated by A. TYBULEWICZ

REFERENCES

1. LIFSHITS E. M. *J. Phys. Chem. Solids.*
2(a). SHUBNIKOV W. and DE HAAS W. *Proc. Acad. Sci., Amst.* **33**, 418 (1930).
2(b). DE HAAS W. and SHUBNIKOV W. *Leid. Commun.* **207**, 210 (1930).
3. LIFSHITS E. M. and KOSEVICH A. *Zh. éksp. theor. fiz.* **29**, 730 (1955).
4. LIFSHITS E. M., AZBEL' M. and KAGANOV M. *Zh. éksp. theor. fiz.* **31**, 64 (1956).
5. LIFSHITS E. M. *Zh. éksp. theor. fiz.* **30**, 814 (1956).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 11–18.

QUANTUM THEORY OF THE ELECTRICAL CONDUCTIVITY OF METALS IN A MAGNETIC FIELD

E. M. LIFSHITS

Physical-Technical Institute, Academy of Science of the Ukranian S.S.R., Kark'off, U.S.S.R.

(*Received* 26 *December* 1956)

Abstract—The paper gives a consistent quantum-mechanical theory of the electrical conductivity of metals in magnetic fields. A kinetic equation is deduced for the statistical operator with general assumptions on the energy spectrum of electrons, and the limit of the transition to a classical case is analysed.

Asymptotes of the kinetic coefficients in strong magnetic fields are studied. Starting from the established relationship between the solutions of quantum and classical equations, quantum corrections to the kinetic coefficients, which include resistance and conduction oscillations, are determined.*

VOL. 4 1958

1. INTRODUCTION

IT is well known that the correct interpretation of electrical conduction in metals is basically quantum in character. Nevertheless, since the electron wavelength is small compared with its free path, and, in the presence of the magnetic field, also compared with the radius of curvature of its trajectory, it is usually sufficient to employ only a semiclassical treatment. This means that the electron may be regarded as a classical quasi-particle, which obeys Fermi statistics with a dispersion law

$$\epsilon = \epsilon_j(p_x, p_y, p_z) \tag{1}$$

where $\epsilon_j(\boldsymbol{p})$ = a periodic function with the period of the reciprocal lattice, j = band number.

On this basis a semiclassical theory of galvanomagnetic effects in metals was developed in.[1]

In such a treatment, however, certain specifically quantum effects (e.g. resistance oscillations in a magnetic field), connected with the quantization of levels, are lost. These effects were found experimentally[2] but their theoretical treatment[3] does not seem to be satisfactory or consistent.

The present treatment aims to develop a consistent quantum theory of metallic conduction in magnetic fields. The theory enables us to study the quantum oscillation effects mentioned above and to find the transition to the classical description.

As shown in [4] the energy levels ϵ_n of a particle, which obeys the law of dispersion (1), in a magnetic field H directed along the z-axis are determined in the quasiclassical approximation by the conduction

$$S(\epsilon_n, p_z) = \frac{e}{c} Hh(n+\gamma) \quad \begin{matrix} (n \gg 1) \\ (0 < \gamma < 1) \end{matrix} \tag{2}$$

where $S(\epsilon, p_z)$ = area of a section of the surface (1) by a plane p_z = const. in quasimomentum space.

The separation of the levels, according to (2), is:

$$\Delta\epsilon_n = \epsilon_{n+1} - \epsilon_n = \hbar\omega^* = \mu^* H \qquad \mu^* = \frac{e\hbar}{m^* c} \tag{3}$$

$$\omega^* = \frac{eH}{m^* c}; \qquad m^* = \frac{1}{2\pi} \cdot \frac{\partial S}{\partial \epsilon}; \qquad m^* = m^*(\epsilon, p_z). \tag{4}$$

We can expect some intrinsic effects at $\hbar\omega^* \gtrsim kT$. On the other hand to use the quasiclassical approximation and the dispersion law (1) we must have $\hbar\omega \ll \xi$ (ξ = chemical potential). We shall take the latter condition to be fulfilled in this treatment.

* Some of the results obtained were briefly reported on in reference (5).

2. KINETIC EQUATION OF THE DENSITY MATRIX

If the electron is in a uniform magnetic field $\boldsymbol{H}$ and a weak electric field $\boldsymbol{E}$, we can represent its Hamiltonian as

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{U} \qquad \hat{U} = -e\boldsymbol{E}\hat{\boldsymbol{r}} \tag{5}$$

where $\hat{\mathcal{H}}_0$ is the kinetic energy operator of the electron in a magnetic field. This operator, in the limiting case of the quasiclassical approximation, is obtained from $\epsilon(\boldsymbol{p})$ by replacing components of the momentum $\boldsymbol{p}$ by operators of the kinetic momentum $\hat{\boldsymbol{p}}$, and by appropriate symmetrization the exact meaning of which is not important in our case.

Since the kinetic momentum operators $\hat{p}_x$, $\hat{p}_y$, $\hat{p}_z$ do not commute in a magnetic field $H_z = H$

$$([\hat{p}_x, \hat{p}_y] = \frac{e}{c}H),$$

we shall describe the electron gas state not by a distribution function, but by a density matrix $\hat{f}$.

In the absence of the electric field ($\hat{U} = 0$), the operator $\hat{f} = \hat{f}^0$ is equal to the equilibrium Fermi function:

$$\hat{f}^0 = f^0(\hat{\mathcal{H}}_0); \qquad f^0(\epsilon) = [1 + e^{(\epsilon-\xi)/\theta}]^{-1} \tag{6}$$

where ξ is the chemical potential, $\theta = kT$.

In the electric field we take

$$\hat{f} = \hat{f}^0 + \hat{f}^1. \tag{7}$$

We shall use quantum numbers n and p_z of the operator $\hat{\mathcal{H}}_0$, and, as shown later, in the quasiclassical approximation it will be sufficient to regard $\hat{f}^1$ as a matrix with respect to n and as a number with respect to p_z. This is because momentum p_z is not quantized.

When several electron bands are considered the band number j enters $\hat{f}$ as a parameter. This follows from the fact that in the classical particle motion j remains constant. In the quasiclassical approximation j can change only on collision, and this may be taken into account in the collision integral. The same applies to various spin states s which are taken care of by introducing an additional parameter s in $\hat{f}$ since the energy levels depend on s:

$$\epsilon_{n,s} = \epsilon_n(p_z) + s\mu_0 H; \qquad s = \pm 1;$$

$$\mu_0 = \frac{e\hbar}{2m_0 c}.$$

Later, unless it will be unavoidable, we shall omit the indices j, s.

The kinetic equation for $\hat{f}^1$ may be written as

$$[\hat{f}^0, \hat{\mathcal{H}}] + [\hat{f}^1, \hat{\mathcal{H}}] + \frac{1}{t_0}\tilde{W}\hat{f}^1 = 0. \tag{8}$$

Here $\tilde{W}/t_0$ denotes a certain linear transformation corresponding to the collision integral, and t_0 is a relaxation time. To make our discussion clearer we shall start from the simplest case $\tilde{W} = 1$.

Since clearly $[\hat{f}^0, \hat{\mathcal{H}}_0] = 0$, and $[\hat{f}^1, \hat{U}]$ which is a quadratic function of the electric field, can be neglected, we have

$$[\hat{f}^1, \hat{\mathcal{H}}_0] + \frac{1}{t_0}\hat{f}^1 = -[\hat{f}^0, \hat{U}]. \tag{9}$$

Taking into account the diagonality of $\hat{\mathcal{H}}_0$ and $\hat{f}^0$, we have

$$\frac{i}{\hbar}(\epsilon_{n'} - \epsilon_n) f_{nn'}{}^1 + \frac{f_{nn'}{}^1}{t_0} = (f_{n'}{}^0 - f_n{}^0)\frac{i}{\hbar}U_{nn'}. \tag{10}$$

From the equations of motion of the electron in a magnetic field it follows that

$$-[\hat{\mathcal{H}}_0, \hat{\boldsymbol{r}}] = \dot{\hat{\boldsymbol{r}}} = \hat{\boldsymbol{V}} \tag{11}$$

or, from (5)

$$\frac{i}{\hbar}(\epsilon_n - \epsilon_{n'})U_{n\boldsymbol{n}'} = e\boldsymbol{E}\hat{\boldsymbol{V}}_{nn'} \tag{11a}$$

where $\hat{\boldsymbol{V}}$ = velocity operator corresponding to the classical quantity $\boldsymbol{V} = \partial\epsilon/\partial\boldsymbol{p}$.

This becomes

$$\left\{\frac{i}{\hbar}(\epsilon_{n'} - \epsilon_n) + \frac{1}{t_0}\right\} f_{nn'}{}^1 = -\frac{f_{n'}{}^0 - f_n{}^0}{\epsilon_{n'} - \epsilon_n} e\boldsymbol{E}\hat{\boldsymbol{V}}_{nn'}. \tag{12}$$

According to (3)

$$\epsilon_{n'} - \epsilon_n = (n' - n)\hbar\omega^*$$

Assuming

$$\hat{f}^1 = -et_0\boldsymbol{E}\hat{\Psi} \tag{13}$$

and reducing (13) by ω^*

$$[i(n'-n)+\gamma]\hat{\Psi}_{nn'} = \gamma g_{nn'}\hat{V}_{nn'}$$

$$\gamma = \frac{1}{t_0\omega^*} = \frac{H_0^*}{H}; \qquad H_0^* = \frac{m^*c}{et_0}; \tag{14}$$

$$g_{nn'} = \frac{f_{n'}^0 - f_n^0}{\epsilon_{n'} - \epsilon_n}.$$

Assuming $n'-n = k$ we obtain finally

$$(ik+\gamma)\Psi_k = \gamma g_k V_k \tag{15}$$

Here

$$\psi_{n,n+k} = \psi_k(\epsilon_n, p_z); \quad V_{n,n+k} = V_k(\epsilon_n, p_z) \tag{16}$$

$$g_k(\epsilon_n) = g_{n,n+k} = g_{n+k,n} = \frac{f^0(\epsilon_n + k\hbar\omega^*) - f^0(\epsilon_n)}{k\hbar\omega^*}. \tag{17}$$

At $k = 0$ the expression g_k is undetermined. A simple calculation shows that†

$$g_0(\epsilon) = \frac{\partial f^0}{\partial \epsilon} \tag{17a}$$

which corresponds to (17a).

For $\hbar\omega^* \to 0$

$$g_k(\epsilon) \to \frac{\partial f^0(\epsilon)}{\partial \epsilon} \tag{18}$$

and (15) assumes the form

$$(ik+\gamma)\psi_k = \frac{\partial f^0}{\partial \epsilon} V_k$$

$$\psi_k = \psi_k(\epsilon, p_z); \qquad V_k = V_k(\epsilon, p_z). \tag{19}$$

† From $\dot{\hat{\mathcal{H}}}_0 = e\boldsymbol{E}\hat{\boldsymbol{V}}$ we find

$$[\dot{f}^0, \hat{U}] = \dot{f}^0 = \lim \frac{f^0(\hat{\mathcal{H}}_0 + e\boldsymbol{E}\boldsymbol{V}\delta t) - f^0(\hat{\mathcal{H}}_0)}{\delta t}.$$

Since for any degree m, due to diagonality of $\hat{\mathcal{H}}_0$, we have

$$\{(\hat{\mathcal{H}}_0 + e\boldsymbol{E}\boldsymbol{V}\delta t)^m - \hat{\mathcal{H}}_0{}^m\} \Big/ \delta t \Big|_{nn} = e\boldsymbol{E}\boldsymbol{V}_{nn} m\epsilon_n{}^{m-1}$$

then for an arbitrary function $f^0(\hat{\mathcal{H}}_0)$ we obtain

$$[\dot{f}^0, \hat{U}] = e\boldsymbol{E}\boldsymbol{V}_{nn} \frac{\partial f^0}{\partial \epsilon_n}$$

This is identical with the classical kinetic equation of the Fourier amplitudes of the distribution function $\psi = \psi(\epsilon, p_z, \varphi)$ (φ = angular variable of the particle trajectory in momentum space, proportional to the period of revolution) given in.[1] V_k has the meaning of the Fourier amplitudes of the classical velocity vector.

$$\boldsymbol{V}(\epsilon, p_z, \varphi) = \sum_k \boldsymbol{V}_k e^{ik\varphi} \qquad \boldsymbol{V} = \frac{\partial \epsilon}{\partial \boldsymbol{p}}$$

$$\boldsymbol{V}_k = \frac{1}{2\pi}\int_0^{2\pi} \frac{\partial \epsilon}{\partial \boldsymbol{p}} e^{ik\varphi}\, d\varphi \qquad \varphi = 2\pi \int \frac{dl}{v_\perp} \Big/ \oint \frac{dl}{v_\perp} \tag{20}$$

$$dl = |d\boldsymbol{p}|; \qquad v_\perp = \sqrt{v_x^2 + v_y^2}. \tag{21}$$

Returning to the quantum equation (15) we shall replace $\boldsymbol{V}_k(\epsilon_n, p_z)$ by their limiting values (21) (this corresponds to the quasiclassical approximation for matrix elements $\boldsymbol{V}_{nn'}$). For g_k a similar replacement ($g_k \to \delta f^0/\delta\epsilon$) is not allowed, since $f^0(\epsilon)$ changes very rapidly near the limiting energy ($\epsilon - \xi \sim \theta$) and the whole sought-after effect is bound up with this.

The solution of the kinetic equation (15) is

$$\psi = \gamma \frac{g_k V_k}{ik+\gamma}. \tag{22}$$

We shall now return to the general kinetic equation (8) and we shall assume that

$$(\tilde{W}\hat{\psi})_{n,n+k} = \hat{W}_{kk'}\psi_{k'} \tag{23}$$

where $\hat{W}_{kk'}$ = operator acting on the variable ϵ_n and p_z in the function $\psi_k(\epsilon_n, p_z)$

$$\hat{W}_{kk'}\psi_{k'} = \Sigma \int w_{kk'}(\epsilon, p_z; \epsilon', p'_z) \times$$

$$\times \psi_{k'}(\epsilon', p'_z)\, dp'_z \Delta\epsilon'. \tag{24}$$

$$(\epsilon = \epsilon_n; \quad \epsilon' = \epsilon_{n'})$$

(If the transitions between various bands are considered, the kernel $W_{kk'}$ has also the indexes j, j'.)

Then, repeating the previous treatment, we transform (8) into a form similar to (15)

$$ik\psi_k + \gamma\hat{W}_{kk'}\psi_{k'} = \gamma g_k V_k. \tag{25}$$

Equations (25) and (15) transform into the

classical equation (missing in the original) of [1] on replacement of $g_k(\epsilon)$ by the limiting expression $\partial f^0/\partial\epsilon$ and on transition from summation in (24) to integration. The quantities $\hat{W}_{kk'}$ in (25), in agreement with the correspondence principle, represent the Fourier amplitudes of the collision operator in equation (33) of reference (1).

Since g_k, and consequently $\hat{\psi}_k$ change rapidly as a function of ϵ_m and slowly as a function of p_z (for fixed $\epsilon_n = \epsilon$), it is convenient to integrate with respect to ϵ and sum with respect to p_z in equations (25) and (24), i.e. ϵ and p_z are regarded as independent variables. Solving (2) $n = n(\epsilon, p_z)$ for p_z we find

$$\sum_n \int F(\epsilon_n, p_z)\, dp_z \Delta\epsilon_n$$

$$= \int \{ \sum_{p_{zn}} F(\epsilon, p_{zn}) \Delta p_{zn} \}\, d\epsilon \qquad (26)$$

$$p_{zn} = p_z(n, \epsilon); \quad \Delta p_{zn} = \frac{\partial p_z}{\partial n} = \frac{ehH}{c} \Big/ \frac{\partial S}{\partial p_z}.$$

(The limits of summation in n now depend on ϵ!) Thus for elastic scattering on impurities

$$\hat{W}_{kk'}\hat{\psi}_{k'} = \sum_{\epsilon'} \int W_{kk'}^{p_z p'_z}(\epsilon)\, \delta(\epsilon - \epsilon') \times$$

$$\times [\psi_k(\epsilon, p_z) - \psi_{k'}(\epsilon', p'_z)]\, dp'_z \Delta\epsilon' \qquad (27)$$

$$= \sum_{p'_z} W_{kk'}^{p_z p'_z}(\epsilon)[\psi_k(\epsilon, p_z) - \psi_{k'}(\epsilon, p'_z)]\Delta p'_z.$$

The quantities $\hat{\psi}_k$, V_k, g_k in (25) and (26) are understood to be functions of a continuous parameter ϵ and a discrete parameter $p_{zn} = p_z(n, \epsilon)$

$$\psi_k = \psi_k(\epsilon, p_{zn}); \; V_k = V_k(\epsilon, p_{zn})$$

$$g_k = g_k(\epsilon, p_{zn}) = \frac{f^0(\epsilon + k\hbar\omega^*) - f^0(\epsilon)}{k\hbar\omega^*}; \qquad (28)$$

$$g_0 = \frac{\partial f^0}{\partial \epsilon}.$$

In scattering on impurities the parameter $\hbar\omega^*/kT$ does not occur in the kernel of the collision operator (27) and therefore the kernel $W_{kk'}^{p_z p'_z}(\epsilon)$ may be replaced by its classical value.

In the nonelastic scattering by phonons such a replacement is not permissible.

As in the classical case, it is not possible to obtain an explicit solution of (25) for arbitrary fields without special assumptions about the collision integral. We shall study therefore the asymptotic behaviour of the electrical conductivity σ_{ik} in large fields ($\gamma \ll 1$) and try to elucidate the connection between quantum and classical expressions for σ_{ik} for arbitrary values of γ. In the latter case the solution of the classical problem will be regarded as known and the quantum effects described in terms of this solution.

3. ASYMPTOTIC TREATMENT OF THE CONDUCTIVITY IN STRONG FIELDS ($\gamma \ll 1$)

As shown in [1] the asymptotic treatment of the conductivity in strong fields is basically determined by the topology of the Fermi surface. Since quantization of the levels takes place only in the case of closed orbits (finite motion), the whole of our quantum discussion applies only to that case.

This means, in particular, that, from the equation of motion in a magnetic field

$$[\hat{\boldsymbol{p}}, \hat{\mathcal{H}}_0] = \frac{e}{c} V \times \boldsymbol{H} \qquad (29)$$

which in terms of matrix elements has the form

$$ik\boldsymbol{p}_k = m^* \boldsymbol{V}_k \times \boldsymbol{n} \qquad \boldsymbol{H} = H\boldsymbol{n} \qquad (30)$$

we find that the diagonal matrix elements of velocity in the plane xy, which is perpendicular to the direction $\boldsymbol{n}$ of the magnetic field, are equal to zero

$$\boldsymbol{V}_{nn} \times \boldsymbol{n} = \boldsymbol{V}_0 \times \boldsymbol{n} = 0. \qquad (31)$$

We shall introduce components of the operator $\hat{\boldsymbol{\psi}}$

$$\hat{\boldsymbol{\psi}} \times \boldsymbol{n} = \hat{\boldsymbol{\varphi}}; \qquad (\hat{\boldsymbol{\psi}}, \boldsymbol{n}) = \hat{\vartheta}. \qquad (32)$$

According to (13)

$$\hat{f}^1 = -et_0\{(\hat{\boldsymbol{\varphi}}, \boldsymbol{E} \times \boldsymbol{n}) + \hat{\vartheta}(\boldsymbol{E}, \boldsymbol{n})\}. \qquad (33)$$

Taking (30) into account, we find from (25)

$$ik\boldsymbol{\varphi}_k + \gamma \hat{W}_{kk'}\boldsymbol{\varphi}_{k'} = ik\gamma \frac{g_k \boldsymbol{p}_k}{m^*}. \qquad (34)$$

Writing down from (34) $\boldsymbol{\varphi}_k$ in powers of γ we shall initially exclude φ_0. Substituting $k = 0$ in (34) we obtain

$$\boldsymbol{\varphi}_0 = -\hat{W}_{00}^{-1}\hat{W}_{0k'}\boldsymbol{\varphi}_{k'} \qquad k' \neq 0. \quad (35)$$

Therefore

$$\boldsymbol{\varphi}_k - \gamma L_{kk'}\boldsymbol{\varphi}_{k'} = \gamma\frac{g_k\boldsymbol{p}_k}{m^*}; \qquad k, k' \neq 0. \quad (36)$$

$$\hat{L}_{kk'} = \frac{i}{k}\{\dot{W}_{kk'} - \dot{W}_{k0}\dot{W}_{00}^{-1}\dot{W}_{0k'}\}$$

$$k, k' \neq 0. \quad (37)$$

and the solution $\boldsymbol{\varphi}_k$ may be easily expressed as a series in powers of γ

$$\boldsymbol{\varphi}_k = (1-\gamma\hat{L})_{kk'}^{-1}\gamma\frac{g_{k'}\boldsymbol{p}_{k'}}{m^*}$$

$$= \gamma\frac{g_k\boldsymbol{p}_k}{m^*} + \gamma\hat{L}_{kk'}\gamma\frac{g_{k'}\boldsymbol{p}_{k'}}{m^*} + \ldots. \quad (38)$$

Similarly, for ϑ_k we have

$$\vartheta_0 = \dot{W}_{00}^{-1}\left\{\frac{\partial f^0}{\partial\epsilon}V_{0z} - \dot{W}_{0k'}\vartheta_{k'}\right\}. \quad (39)$$

Hence

$$\vartheta_k - \gamma\hat{L}_{kk'}\vartheta_{k'} = \gamma h_k;$$

$$\gamma h_k = \frac{\gamma}{ik}\left\{g_k V_k - \dot{W}_{k0}\dot{W}_{00}^{-1}\frac{\partial f^0}{\partial\epsilon}V_{0z}\right\} \quad (40)$$

and consequently

$$\vartheta_k = \gamma h_k + \gamma\hat{L}_{kk'}\gamma h_k + \ldots \quad (41)$$

$$\vartheta_0 = \dot{W}_{00}^{-1}\frac{\partial f^0}{\partial\epsilon}V_{0z} - \dot{W}_{00}^{-1}W_{0k'}\gamma h_{k'} + \ldots. \quad (41a)$$

We see that, as in the classical case, all elements of $\boldsymbol{\varphi}_k$ and ϑ_k, with the exception of ϑ_0, begin with the first degree of $\gamma(\sim 1/H)$. Also the terms in $\boldsymbol{\varphi}_k\,(k \neq 0)$ which are linear in γ do not depend on the collision integral.

The current is given by

$$\boldsymbol{j} = e\mathcal{S}p\hat{f}^1\hat{\boldsymbol{V}}. \quad (42)$$

Thus the terms linear in $\gamma\sim 1/H$ which occur in the components σ^{xz}, σ^{zy}, σ^{yy} of the conductivity tensor, are not related to the collision integral and are easily obtained in explicit form. To find the components j^x and j^y we shall calculate $\boldsymbol{j}\times\boldsymbol{n}$. Using (29) and (8)

$$\boldsymbol{j}\times\boldsymbol{n} = e\mathcal{S}p\hat{f}^1\hat{\boldsymbol{V}}\times\boldsymbol{n}$$

$$= \frac{c}{H}\mathcal{S}p\hat{f}^1[\hat{\boldsymbol{p}}, \hat{\mathcal{H}}_0]$$

$$= \frac{c}{H}\mathcal{S}p\hat{\boldsymbol{p}}[\hat{\mathcal{H}}_0\hat{f}^1]$$

$$= \frac{c}{H}\mathcal{S}p\hat{\boldsymbol{p}}[\hat{f}^0\dot{U}] + \frac{c}{Ht_0}\mathcal{S}p\hat{\boldsymbol{p}}\tilde{W}\hat{f}^1. \quad (43)$$

From (33) and (38), retaining the terms $\sim 1/H$ we find

$$\frac{c}{Ht_0}\mathcal{S}p\hat{\boldsymbol{p}}\tilde{W}\hat{f}^1 = \frac{ec}{H}E_z\mathcal{S}p\hat{\boldsymbol{p}}\tilde{W}\hat{\vartheta}.$$

The terms proportional to E_x and E_y are given by the first component. Since $\hat{U} = -eE_iX^i$, therefore

$$\boldsymbol{j}\times\boldsymbol{n} = \frac{ec}{H}E_i\mathcal{S}p\hat{\boldsymbol{p}}[\hat{f}^0, \dot{X}^i]. \quad (44)$$

Considering the current components j^x and j^y we find, with an accuracy to $\sim 1/H^2$,

$$\sigma_{ik} = \sigma_{ik}^{(1)} + 0\left(\frac{1}{H^2}\right)$$

$$\sigma_{xx}^{(1)} = -\frac{ec}{H}\mathcal{S}p\hat{p}_y[\hat{f}^0, \hat{x}]$$

$$\sigma_{yy}^{(1)} = \frac{ec}{H}\mathcal{S}p\hat{p}_x[\hat{f}^0, \hat{y}]$$

$$\sigma_{xy}^{(1)} = \frac{ec}{H}\mathcal{S}p\hat{p}_x[\hat{f}^0, \hat{x}]. \quad (45)$$

From the equality

$$\mathcal{S}p\hat{p}_y[\hat{f}^0, \hat{x}] = \mathcal{S}pf^0[\hat{x}, \hat{p}_y] = 0$$

and a similar equality for $\sigma_{yy}^{(1)}$ we find

$$\sigma_{xx}^{(1)} = \sigma_{yy}^{(1)} = 0. \quad (46)$$

Since $[\hat{p}_x, \hat{x}] = 1$ the transformation of the expression for $\sigma_{xy}^{(1)}$ should be carried out with caution (direct change of the order of the multipliers in $\mathcal{S}p\ldots$ is not permitted). Detailed calculations show that when all the isoenergetic surfaces (1) in the interval $|\epsilon-\zeta| \sim \theta$ are closed then for each separate band

$$\mathcal{S}p\hat{p}_x[\hat{f}^0, \hat{x}] = \mathcal{S}pf^0 = N_+$$

when $\partial\epsilon/\partial n > 0$

$$\mathcal{S}p\hat{p}_x[\hat{f}^0, \hat{x}] = \mathcal{S}p(\hat{f}^0 - 1) = -N_- \tag{47}$$

when $\partial\epsilon/\partial n < 0$

where $N_+ =$ number of occupied states with $m^* > 0$ (number of "electrons") and $N_- =$ number of unoccupied states with $m^* < 0$ (number of "holes"). In this case

$$\sigma_{xy}^{(1)} = \frac{ec}{H}(\Sigma N_+ - \Sigma N_-) \tag{48}$$

which is identical with the classical result obtained earlier. For a more complex topology of the Fermi surface, just as in the classical case, equation (48) no longer holds.

When the number of the "electrons" and the "holes" is the same

$$\sigma_{xy}^{(1)} = 0. \tag{49}$$

We have shown now that those terms of σ_{ik} (written down as a series in γ) which become zero in the classical case need no quantum corrections.

This means that in the limit of the quasiclassical approximation special effects may be expected in the temperature dependence of conductivity and in the quantum oscillations. We shall now consider these questions.

4. RELATIONSHIP BETWEEN QUANTUM AND CLASSICAL EXPRESSIONS FOR THE CONDUCTIVITY IN A MAGNETIC FIELD

The conductivity tensor, according to (42) and (13) has the form $\sigma^{++'}$

$$\sigma^{xx'} = -e^2 t_0 \mathcal{S}p\hat{V}^x\hat{\psi}^{x'}. \tag{50}$$

Considering $\hat{V}$ and $\hat{\psi}$ and their diagonality with respect to p_z we find

$$\sigma^{xx'} = A\int \sum_{n,n'} V_{nn'}{}^x \psi_{n'n}{}^{x'}\, dp_z$$

$$= A\int \{\sum_n V_{nn}{}^x \psi_{nn}{}^{x'} + 2\mathrm{Re}\sum_{n'>n} V_{n'n}{}^x \psi_{nn'}{}^{x'}\}\, dp_z$$

$$= A\sum_n \int \{V_0{}^x \psi_0{}^{x'} + 2\mathrm{Re}\sum_{k=1}^{\infty} V_{-k}{}^x \psi_k{}^{x'}\}\, dp_z. \tag{51}$$

A normalized constant A, can be simply found by comparison with the limiting classical case $\hbar\omega \to 0$ (see p. 75 of reference (1).

Since $m^*\Delta\epsilon_n = eH\hbar/c$, therefore

$$A = -\frac{e^2 t_0}{\hbar^3}\cdot\frac{2eHh}{c}. \tag{52}$$

For $\psi_k{}^x$ in (51) we should substitute the solution of the kinetic equation (25). Since (25) and (51) are linear with respect to $\psi_k{}^x$ and g_k results of such a substitution may be represented as a "scaler product"

$$\sigma^{xx'} = -\frac{e\hbar H}{c} 2(\chi_k^{xx'}, g)$$

$$(\chi^{xx'}, g) = \sum_n \int \{\sum_{k=0}^{\infty} \chi_k^{xx'} g_k\}\, dp_z;$$

$$\chi = \chi_k(\epsilon, p_z). \tag{53}$$

Since $V_0{}^x V_0{}^y = 0$ and g_0 is present only in the equation for ψ^z, then none of the components of $\sigma^{xx'}$, except σ^{zz}, contains $g_0 = \partial f_0/\partial\epsilon$.

$$\sigma^{xx'} = -2\int \sum_{pzn} \{\sum_{k=1}^{\infty} \chi_k^{xx'} g_k\} \times$$

$$\times m^*\Delta p_{zn}\, d\epsilon; \qquad x, x' \neq z, z \tag{54}$$

$$\sigma^{zz} = -2\int \sum_{pzn} \left\{ \chi_0{}^{zz}\frac{\partial f^0}{\partial\epsilon} + \sum_{k=1}^{\infty} \chi_k^{zz} g_k \right\} m^*\Delta p_{zn}\, d\epsilon. \tag{55}$$

When $\tilde{W} = 1$, according to (22) we have

$$\chi_k^{xx'} = \frac{2\pi e^2 t_0}{\hbar^3}\mathrm{Re}\frac{2\gamma}{ik+\gamma} V_k{}^x V_k{}^{x'}. \tag{56}$$

For the diagonal components

$$\chi_k^{xx} = \frac{4\pi e^2 t_0}{\hbar^3}\frac{\gamma^2}{k^2+\gamma^2}|V_k{}^x|^2. \tag{57}$$

The expressions (56) and (57) are identical with the classical equations since, as shown above V_k is the Fourier amplitude of the classical velocity vector; to find the classical expression for $\sigma^{xx'}$ it is sufficient to

replace g_k by $\partial f_0/\partial\varepsilon$ in (54) and (55) and instead of summing n to integrate with respect to $d\varepsilon$.

In the general case the quantities $\chi^{xx'}$ may be expressed in terms of a "Green's function" of the kinetic equation (25):

$$ik\mathcal{J}_k+\gamma W_{kk'}\mathcal{J}_{k'}$$

$$=\gamma\delta(\epsilon-\epsilon^0)\frac{\delta p_{zn},p_{zn'}}{\Delta P_{zn}}\delta_{kk^0} \tag{58}$$

$$\mathcal{J}_k=\mathcal{J}_{\epsilon p_z k}^{\epsilon^0,\,p_z{}^0,\,k^0};$$

$$p_z=p_{zn}=p_z(n,\epsilon); \qquad p_z{}^0=p_{zn^0}$$

by means of the equality

$$m^*(\epsilon^0,p_z{}^0)\chi_{k^0}{}^{xx'}(\epsilon^0,p_z{}^0)$$

$$=\frac{2\pi e^2 t_0}{\hbar^3}2\mathrm{Re}\int\sum_{pz}\{\sum_k V_k{}^{x'}(\epsilon^0,p_z{}^0)\,.\,\mathcal{J}_{\epsilon p_z k}^{\epsilon^0 p_z{}^0,\,k^0}\cdot V_{-k}{}^x(\epsilon,p_z)\}m^*(\epsilon,p_z)\,d\epsilon dp_z. \tag{59}$$

The classical "Green's function" is the solution of the equation

$$ik\mathcal{J}_k+\gamma W_{kk'}\mathcal{J}_{k'}$$
$$=\gamma\delta(\epsilon-\epsilon^0)\delta(p_z-p_z{}^0)\delta_{kk^0} \tag{60}$$

where $W_{kk'}$ = the classical limit of the collision integral; to find $\chi^{xx'}$ integration with respect to p_z should replace summation in (59).

For the case when only the elastic scattering on impurities is important, the collision integral does not contain temperature. Consequently, the quantity $\chi^{xx'}$ does not contain the parameter $\mu^*H/\theta=\hbar\omega^*/\theta$ and the transition to the classical limit in $\chi^{xx'}$ involves omission of the quantities $\sim\hbar\omega^*/\xi$, compared with $\hbar\omega/\theta$, which enter the quantum effects through the expressions

$$g_k=\frac{f^0(\epsilon+k\hbar\omega^*)-f^0(\epsilon)}{k\hbar\omega^*}.$$

Consequently in the above case we can regard $\chi_k{}^{xx}$ in (54) and (55) to be equal to their classical analogues.

The classical expressions for the conductivity are found by the replacement $g_k\to\partial f^0/\partial\varepsilon$ and integration

$$\sigma^{xx'}=-2\int\int\frac{\partial f^0}{\partial\epsilon}\chi^{xx'}m^*\,dp_z d\epsilon$$

$$x,x',\neq z,z$$

$$\sigma^{zz}=-2\int\int\frac{\partial f^0}{\partial\epsilon}(\chi_0{}^{zz}+\chi^{zz})m^*\,dp_z d\epsilon \tag{61}$$

$$\chi^{xx'}=\sum_{k=1}^{\infty}\chi_k{}^{xx'}.$$

Since the main part of the quantum corrections is due to the neighbourhood of the points of pronounced instability of the function f_0 (i.e. $\varepsilon_n\cong\xi$) and the functions $m^*(\varepsilon,p_z)$ and $\chi(\varepsilon,p_z)$ are continuously variable, we find, from (54) and (17), that the difference $\sigma^{xx'}-\sigma^{xx'}{}_{\mathrm{class}}$ (at $x,x'\neq z,z'$) is given with the same accuracy by

$$\sigma^{xx'}-\sigma_{\mathrm{class}}{}^{xx'}=-2\Bigg\{\sum_{k=1}^{\infty}\int\sum_n\left\{\chi_k^{xx'}\frac{f^0(\epsilon_n+k\hbar\omega^*)-f^0(\epsilon_n)}{k\hbar\omega^*}\right\}\Delta\epsilon_n\,dp_z-$$

$$-\sum_{k=1}^{\infty}\int\int\chi_k^{xx'}\frac{\partial f^0}{\partial\epsilon}m^*\,d\epsilon dp_z\Bigg\}$$

$$=-2\Bigg\{\sum_{k=1}^{\infty}\int dp_z\Bigg\{\sum_{n\geqslant k}m^*\chi_k^{xx'}(\epsilon_n-k\hbar\omega^*)\frac{f^0n}{k}-\sum_{n\geqslant 0}m^*\chi_k^{xx'}(\epsilon_n)\frac{f^0n}{k}-$$

$$-f^0(\epsilon_0)m^*(\epsilon_0,p_z)\chi_k^{xx'}(\epsilon_0,p_z)+\int f^0\frac{\partial(\chi_k^{xx'}m^*)}{\partial\epsilon}d\epsilon\Bigg\}$$

$$= 2\sum_{k=1}^{\infty}\int dp_z\left\{\sum_{\epsilon_n}\frac{\partial(\chi_k^{xx'}m^*)}{\partial\epsilon}f^0(\epsilon_n)\Delta\epsilon_n - \int\frac{\partial(\chi_k^{xx'}m^*)}{\partial\epsilon}f^0\,d\epsilon\right\}.$$

Summing over k we find that

$$\sigma^{xx'} - \sigma_{\text{class}}{}^{xx'} = \left\{\int\sum_n f^0\frac{\partial(\chi^{xx'}m^*)}{\partial\epsilon}\Delta\epsilon\,dp_z - \iint f^0\frac{\partial(\chi^{xx'}m^*)}{\partial\epsilon}\,d\epsilon dp_z\right\}.$$

Integrating with respect to n, using the equality $\partial\epsilon/\partial n = \hbar\omega^* = \Delta\epsilon$ and the Poisson formula we obtain

$$\sigma^{xx'} - \sigma_{\text{class}}{}^{xx'} = 2\left\{\sum_n \phi^{xx'}(n) - \int\phi^{xx'}(n)\,dn\right\} = \phi^{xx'}(0) + \sum I_k{}^{xx'}$$

$$\phi^{xx'}(n) = \int f^0\frac{\partial(\chi^{xx'}m^*}{\partial\epsilon}\Big|_{\epsilon=\epsilon(n,p_z)}\cdot\hbar\omega^*\,dp_z; \tag{62}$$

$$I_k{}^{xx'} = 2\int\phi^{xx'}(n)e^{2\pi ikn}\,dn.$$

The expression for the quantum correction obtained here contains both the oscillating part and the part which varies continuously with the field. In these corrections the levels with small n (in particular ϕ (0)) are important. The quasiclassical approximation is accurate for these levels only when the dispersion law is quadratic. In the general case, however, we shall retain only the oscillating part of the conductivity since the continuously variable part represents only a small correction of the fundamental classical value. Thus

$$\sigma^{xx'} - \sigma_{\text{class}}{}^{xx'} = \sum_{k=1}^{\infty} I_k{}^{xx'};$$

$$I_k{}^{xx'} = 2\int\int f^0\frac{\partial(\chi^{xx'}m^*)}{\partial\epsilon}\hbar\omega^*e^{2\pi ikn}\,dndp_z. \tag{63}$$

According to (55) σ^{zz} contains an additional term with $\partial f^0/\partial\epsilon\;\chi_0{}^{zz}$ and therefore by a procedure similar to that followed above

$$\sigma^{zz} - \sigma_{\text{class}}{}^{zz} = \sum_k (I_k{}^{zz} - L_k{}^{zz}) \tag{64}$$

$$L_k{}^{zz} = 2\int\int\frac{\partial f^0}{\partial\epsilon}\chi_0{}^{zz}m^*\hbar\omega^*e^{2\pi ikn}\,dndp_z.$$

The expressions (63) and (64) are of such form that the oscillating components, contained in the integrals T_k and L_k, may be easily separated out. A detailed analysis of these components and a discussion of various aspects of the Shubnikov–de Haas effect are given in the following paper.

REFERENCES

1. LIFSHITS E. M., ASBEL' M. JA. and KAGANOV M. J. *J. Phys. Chem. Solids.* To be published.
2. DE HAAS *et al. Commun. Phys. Lab. Univ. Leiden* **237** (1935).
3. ACHIESER A. J. *Zh. eksp. theor, fiz* **9,** 726 (1939); DAVIDOV B. J. and POMERANCHUK J. JA. *Zh. eksp. theor. fiz.* **9,** 1295 (1939); ZILBERMAN G. E. *Zh. eksp. theor. fiz.* **25,** 762 (1955).
4. LIFSHITS E. M. and KOSEVITCH A. M. *Zh. eksp. theor. fiz.* **29,** 730 (1955).
5. LIFSHITS E. M. *Zh. eksp. theor. fiz.* **30,** 817 (1956).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 19–26.

QUANTUM THEORY OF LONGITUDINAL MAGNETO-RESISTANCE

P. N. ARGYRES

Westinghouse Research Laboratories, Pittsburgh 35, Pennsylvania

(*Received* 23 *April* 1957)

Abstract—An extension of previous work on the effect of quantization of electron orbits in a magnetic field on the lattice relaxation time to include scattering among different oscillator states has been made. It is found that the relaxation time is inversely proportional to the density of states in the presence of the magnetic field. The current for a longitudinal electric field is calculated and found to depend, in general, on the magnetic field, in contradistinction to the classical theory, and to present, in particular, an oscillatory behavior with the magnetic field under degenerate conditions.

VOL. 4 958

1. INTRODUCTION

IN a recent paper[1]* the effect of a strong magnetic field on the scattering of the conduction electrons in a simple semiconductor was discussed. It was found that in the case of a magnetic field so strong that all electrons occupy the ground oscillator state the scattering by either acoustical lattice vibrations or ionized impurities can be described by a relaxation time that depends on the magnetic field strength. For lattice scattering, for example, the relaxation time was found to be inversely proportional to the magnetic field. The longitudinal electrical conductivity in such a case depends also on the magnetic field. For example, for a non-degenerate semiconductor in the lattice scattering range the magnetoresistance is a linear function of the magnetic field strength, whereas in the ionized-impurity range a negative magnetoresistance is possible.

In this note we wish to consider the effect of a magnetic field of arbitrary strength on the scattering of carriers in a simple semiconductor by acoustical lattice vibrations. The carriers now will be distributed among many oscillator states and will be scattered from one to another whenever it is energetically possible for them. The probability of scattering of carriers increases whenever their kinetic energy along the direction of the magnetic field is equal to the energy of a higher oscillator state, because of the increase of the density of accessible final states. Thus the quantization of the electron orbits in the magnetic field and the resulting density of states will give rise to a relaxation time different from the one calculated in the absence of the magnetic field. One, of course, expects that for vanishing magnetic fields the two relaxation times are identical.

The general idea of the treatment is to make a transport theory using the eigenstates of the carrier in the magnetic field. In this way we avoid two difficulties with the usual Boltzmann transport theory, as far as the magnetic field is concerned. First, we do not treat the magnetic field as a perturbation over the motion of the electrons in the lattice and thus do not restrict ourselves to small magnetic fields. Secondly, we properly take into account the effect of the magnetic field on the scattering of the carriers, something that is left out of account in the usual Boltzmann transport theory.

We assume for reasons of simplicity an electron gas with a single energy surface with isotropic effective mass. We also consider here only the case of parallel electric and magnetic fields. The usual Boltzmann theory predicts zero magnetoresistance for such a system.

The case of transverse electric and magnetic fields presents some interesting points. TITEICA[2]

* Comments and references to previous work on this topic are given in this paper.

and DAVYDOV and POMERANCHUK[3] have carried out calculations for this case. Their reasoning of the calculation of the current density is based, however, on classical concepts, and thus it is not clear that their method is strictly correct. It is obvious that in a representation in which the electron energy in the magnetic field is diagonal the transverse current is associated with the off-diagonal matrix elements of the density operator. It is, therefore, essential to generalize the Boltzmann equation to include all elements of the density matrix. We shall report our attempt in this direction in another publication.

In the following section we recall the appropriate wave functions and energy spectrum for an electron in a magnetic field and calculate the matrix elements of $e^{iq\cdot r}$ in this representation noting some useful properties of them. Next we calculate the momentum transfer relaxation time for acoustic lattice vibrations. Finally, we calculate the conductivity of a degenerate and non-degenerate electron gas for a longitudinal electric field.

2. WAVE FUNCTIONS AND MATRIX ELEMENTS

We consider a one electron problem in a magnetic field H along the z axis. For the gauge $A_x = 0$, $A_y = Hx$, $A_z = 0$ the Hamiltonian is, in the effective mass approximation,

$$\mathcal{H} = \frac{1}{2m}[p_x^2+(p_y+m\omega x)^2+p_z^2], \qquad (2.1)$$

where $\omega = eH/mc$ is the cyclotron frequency and m the effective mass of the electron. The eigenvalues and eigenfunctions are:

$$\epsilon(nk_yk_z) = \hbar\omega(n+1/2)+\hbar^2k_z^2/2m \qquad (2.2)$$

$$\psi(nk_yk_z) = \phi_n(x+\lambda^2k_y)e^{ik_yy}e^{ik_zz}L_y^{-1/2}L_z^{-1/2}. \qquad (2.3)$$

Here k_y and k_z are the wave numbers associated with the y and z co-ordinates and L_y, L_z are the corresponding normalization lengths. $\varphi_n(x+\lambda^2k_y)$ is the normalized wave function for a harmonic oscillator of angular frequency ω in its nth excited state and oscillating about the point

$$-\lambda^2k_y = -\hbar k_y/m\omega, \qquad (2.4)$$

i.e. it satisfies the equations

$$\left[\frac{1}{2m}p_x^2+\frac{m\omega^2}{2}(x+\lambda^2k_y)^2\right]\phi_n(x+\lambda^2k_y) = \hbar\omega(n+\tfrac{1}{2})\phi_n(x+\lambda^2k_y) \qquad (2.5)$$

$$\int_{-\infty}^{+\infty}\phi_n^2(x+\lambda^2k_y)\,dx = 1. \qquad (2.6)$$

The density of orbital states per unit wave number interval with a given oscillator quantum number n is

$$N(n,k_z) = (2\pi\lambda)^{-2} = (2\pi)^{-2}(m\omega/\hbar). \qquad (2.7)$$

Thus, the density of orbital states with total energy between ϵ and $\epsilon+d\epsilon$ is easily found to be

$$N(\epsilon)\,d\epsilon = \frac{1}{(2\pi)^2}\frac{1}{2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\times \hbar\omega\sum_n[\epsilon-(n+\tfrac{1}{2})\hbar\omega]^{-1/2}\,d\epsilon, \qquad (2.8)$$

where the sum is understood to extend over all those non-negative integers n for which $\epsilon-(n+\frac{1}{2})\hbar\omega$ is positive.

The average value of the velocity for the state $\psi(nk_yk_z)$ is found to be

$$\bar{v}_x = \bar{v}_y = 0, \qquad \bar{v}_z = \hbar k_z/m. \qquad (2.9)$$

We shall need in the following the matrix elements of $e^{iq\cdot r}$ between two energy states. We find

$$\langle u'k_y'k_z'|e^{iq\cdot r}|nk_yk_z\rangle = \delta(k_y'-q_y-k_y)\,\delta(k_z'-q_z-k_z)J_{n'n}(q_x,k_y',k_y), \qquad (2.10)$$

where

$$J_{n'n}(q_x,k_y',k_y) = \int_{-\infty}^{+\infty}\phi_{n'}(x+\lambda^2k_y')e^{iq_xx}\phi_n(x+\lambda^2k_y)\,dx. \qquad (2.11)$$

The quantities $J_{n'n}$ have some useful properties[2] which make the treatment of scattering among different oscillator states tractable. First $|J_{n'n}(q_x, q_y+k_y, k_y)|^2$ is a function of $(q_x^2+q_y^2)$ only, as can easily be proved with the use of the generating

function for the φ_n; it is actually a polynominal in $(q_x^2+q_y^2)$ multiplied by

$$\exp[-(\hbar^2/2m)(q_x^2+q_y^2)/\hbar\omega].$$

Second, for all n and n'

$$\int_{-\infty}^{+\infty}\int |J_{n'n}(\pm q_x, \pm q_y + k_y, k_y)|^2 \, dq_x dq_y = 2\pi/\lambda^2. \tag{2.12}$$

This is proved as follows:

$$\int_{-\infty}^{+\infty}\int |J_{n'n}|^2 \, dq_x dq_y = \int_{-\infty}^{+\infty}\int dq_x dq_y \int_{-\infty}^{+\infty} \phi_{n'}[x+\lambda^2(k_y+q_y)] \exp(-iq_x x)\, \phi_n(x+\lambda^2 k_y)\, dx \times$$

$$\int_{-\infty}^{+\infty} \phi_{n'}[y+\lambda^2(k_y+q_y)] \exp(iq_x y)\, \phi_n(y+\lambda^2 k_y)\, dy = \frac{1}{\lambda^2}\int_{-\infty}^{+\infty} ds \int_{-\infty}^{+\infty} \phi_{n'}(x+s)\phi_n(x)\, dx \times$$

$$\int_{-\infty}^{+\infty} \phi_{n'}(y+s)\phi_n(y)\, dy \int_{-\infty}^{+\infty} \exp[iq_x(y-x)]\, dq_x.$$

Since, however,

$$\int_{-\infty}^{+\infty} \exp[iq_x(y-x)]\, dq_x = 2\pi\delta(y-x),$$

we have after integration over y,

$$\int_{-\infty}^{+\infty}\int |J_{n'n}|^2 \, dq_x dq_y =$$

$$(2\pi/\lambda^2)\int_{-\infty}^{+\infty} \phi_n^2(x)\, dx \int_{-\infty}^{+\infty} \phi_{n'}^2(x+s)\, ds = 2\pi/\lambda^2$$

on account of equation (2.6).

3. RELAXATION TIME FOR ACOUSTICAL SCATTERING

For the electron lattice interaction we shall take, as in[1],

$$H_1 = E_1 \operatorname{div} u(\boldsymbol{r}), \tag{3.1}$$

where the function u describes the displacement of the atoms due to the lattice vibrations.

We want to calculate the rate of loss of momentum in the z-direction. We can write

$$\dot{k}_z = \sum_{n'k_y'k_z'} W(nk_yk_z, n'k_y'k_z')(k_z'-k_z), \tag{3.2}$$

where W is the rate of scattering from state $\psi(nk_yk_z)$ to $\psi(n'k_y'k_z')$. We are interested in the case where the electron energy is very small, $\epsilon \sim kT$. Then the energy $\hbar\omega(\boldsymbol{q})$ of the phonons absorbed or emitted with appreciable probability is much smaller than that of most of the electrons. We can, therefore, take the collisions as elastic and the distribution of the involved phonons as classical. We then find (see also reference, (1) equation (3.4)) for the rate of transition due to acoustical lattice vibrations

$$W(\alpha,\alpha') = A\delta(\epsilon_{\alpha'}-\epsilon_\alpha) \times \frac{1}{V}\sum_{\boldsymbol{q}} |\langle\alpha'|e^{i\boldsymbol{q}\cdot\boldsymbol{r}}|\alpha\rangle|^2, \tag{3.3}$$

where

$$A = (2\pi/\hbar)(E_1^2 kT/\rho s^2). \tag{3.4}$$

Here α stands for the quantum numbers n, k_y, k_z, s is the sound velocity, V is the volume and ρ the mass density of the specimen. We have taken $\omega(\boldsymbol{q}) = sq$ for the acoustical lattice vibrations.

Using equation (2.9) for the matrix element we find

$$W(nk_yk_z, n'k_y'k_z') = \delta(\epsilon'-\epsilon)\frac{1}{V}\sum_{q_x} |J_{n'n}(q_x, k_y', k_y)|^2,$$

where $\epsilon' = \epsilon(n'k_z')$ and $\epsilon = \epsilon(nk_z)$. The transition probability to all states with given n' and k_z' is

$$W(nk_yk_z, n'k_z') = \sum_{k_y'} W(nk_yk_z, n'k_y'k_z').$$

Because of equation (2.12) we see that

$$W(nk_yk_z, n'k_z') = A\delta(\epsilon'-\epsilon)\frac{2\pi}{L_z}\frac{1}{(2\pi)^2}\frac{1}{\lambda^2}. \tag{3.5}$$

Thus the transition probabilities to states $\psi(n', k_z')$ and $\psi(n', -k_z')$ are equal to each other and therefore forward and backward scattering in the z-direction are equal.

If we define the relaxation time for an electron in state $\psi(nk_yk_z)$ by $1/\tau = -\dot{k}_z/k_z$ we get from equation (3.2)

$$\frac{1}{\tau} = \sum_{n',k_z'} W(nk_yk_z, n'k_z').$$

With the help of equations (3.5), (2.7) and (2.8) this gives

$$\frac{1}{\tau} = A\int \delta(\epsilon'-\epsilon)N(\epsilon')\,d\epsilon' = AN(\epsilon), \tag{3.6}$$

where $N(\epsilon)$ is the density of orbital states given by equations (2.8). We note that the relaxation time for the state $\psi(nk_yk_z)$ depends only on the energy of that state. Also, for values of ϵ equal to $(n'+1/2)\hbar\omega$ ($n' = 0, 1, 2, 3, \ldots$) the relaxation time becomes zero due to the infinite density of states at these values of energy. The effect of the magnetic field on the lattice relaxation time is directly given by its effect on the density of states, which gives $\tau(\epsilon)$ an oscillatory character of "period" $\hbar\omega$.

Writing equation (3.6) explicitly we have

$$\frac{1}{\tau(\epsilon, H)} = B\hbar\omega \sum_{n'} [\epsilon-(n'+\tfrac{1}{2})\hbar\omega]^{-1/2}, \tag{3.7}$$

where the summation extends over all non-negative integers for which $\epsilon-(n'+1/2)\hbar\omega > 0$, and $B = (m^{3/2}/2^{1/2}\pi\hbar^4)(E_1^2kT/\rho s^2)$. We may write

$$1/\tau(\epsilon_n) = \sum_{n'} 1/\tau_{nn'} \tag{3.8}$$

where $\tau_{nn'}$ is a relaxation time for the scattering of an electron from oscillator state n to state n'.

In the absence of the magnetic field we find in the same manner that the relaxation time $\tau_0(\epsilon)$ of an electron with energy $\epsilon = \hbar^2k^2/2m$ is

$$\frac{1}{\tau_0(\epsilon)} = AN_0(\epsilon) = 2B\epsilon^{1/2}, \tag{3.9}$$

where $N_0(\epsilon)$ is the density of orbital states in the field-free case. A and B are the same constants as before.

It is worth noting that for vanishingly small magnetic field $\tau(\epsilon, H)$ equals $\tau_0(\epsilon)$. For, according to the definition of the integral, we may write

$$\lim_{H\to 0} \hbar\omega \sum_{n'} [\epsilon-(n'+1/2)\hbar\omega]^{-1/2} = \int_0^{\epsilon} (\epsilon-x)^{-1/2}\,dx = 2\epsilon^{1/2}. \tag{3.10}$$

For finite magnetic fields we can use Poisson's summation formula.[4] We find

$$\hbar\omega \sum_{n'} [\epsilon-(n'+1/2)\hbar\omega]^{-1/2} = 2\epsilon^{1/2}+2(\hbar\omega)^{1/2}\sum_{r=1}^{\infty}\frac{(-1)^r}{r^{1/2}}\left[\cos\left(2\pi r\frac{\epsilon}{\hbar\omega}\right)\times C\left(2\sqrt{r\frac{\epsilon}{\hbar\omega}}\right)+\sin\left(2\pi r\frac{\epsilon}{\hbar\omega}\right)S\left(2\sqrt{r\frac{\epsilon}{\hbar\omega}}\right)\right], \tag{3.11}$$

where

$$C(u) = \int_0^u \cos\left(\frac{\pi}{2}x^2\right)dx,$$

$$S(u) = \int_0^u \sin\left(\frac{\pi}{2}x^2\right)dx.$$

For $\epsilon/\hbar\omega \gg 1$ this reduces to

$$\hbar\omega \sum_{n'} [\epsilon-(n'+1/2)\hbar\omega]^{-1/2} \cong 2\epsilon^{1/2}+2^{1/2}(\hbar\omega)^{1/2}\sum_{r=1}^{\infty}\frac{(-1)^r}{r^{1/2}}\cos\left(2\pi r\frac{\epsilon}{\hbar\omega}-\frac{\pi}{4}\right) \tag{3.12}$$

and, therefore

$$\frac{1}{\tau(\epsilon,H)} \cong \frac{1}{\tau_0(\epsilon)} \times \left[1+\left(\frac{\hbar\omega}{2\epsilon}\right)^{1/2} \sum_{r=1}^{\infty} \frac{(-1)^r}{r^{1/2}} \cos\left(2\pi r \frac{\epsilon}{\hbar\omega} - \frac{\pi}{4}\right)\right]. \tag{3.13}$$

For $\epsilon < (3/2)\ \hbar\omega$, i.e. for the ground oscillator state, equation (3.7) gives

$$\frac{1}{\tau(\epsilon)} = B\hbar\omega(\epsilon - \tfrac{1}{2}\hbar\omega)^{-1/2} = B\hbar\omega\epsilon_z^{-1/2}, \tag{3.14}$$

in agreement with equation (3.9) of reference (1).

4. LONGITUDINAL MAGNETORESISTANCE

In this section we will calculate the current induced by an electric field E in the z-direction. Let $f(nk_yk_z)$ be the distribution function referring to the eigenstates $\psi(nk_yk_z)$ of the electron in the magnetic field. Since the motion of the electron in the z-direction is a uniformly accelerated one, the Boltzmann equation of the distribution function for the steady state is

$$-\frac{e}{\hbar}E\frac{\partial f}{\partial k_z} = \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = -\frac{f-f_0(\epsilon)}{\tau(\epsilon)}, \tag{4.1}$$

where $\tau(\epsilon)$ is the relaxation time (3.7), $f_0(\epsilon)$ is the distribution function for thermal equilibrium in the absence of the electric field and $\epsilon = \epsilon(nk_z)$.

To the first order in the electric field

$$f(nk_yk_z) = f_0(\epsilon) + (e\hbar/m)Ek_z\tau(\epsilon)f_0'(\epsilon). \tag{4.2}$$

The current density in the z-direction is then

$$J_z = 2\sum_{n=0}^{\infty} \int_{-\infty}^{+\infty} dk_z N(k_z)(-e\bar{v}_z) f(n,k_z),$$

where $N(k_z)\,dk_z$ is the density of orbital states with k_z in the range k_z, k_z+dk_z and a given oscillator quantum number, and the factor 2 takes into account the spin. According to equations (2.7), (2.9) and (4.2) we have

$$\sigma = \frac{J_z}{E} = -\frac{1}{\pi^2}\frac{e^2}{m} \times \left.\begin{aligned} &\times \hbar\omega \sum_{n=0}^{\infty} \int_0^{\infty} k_z^2 \tau(\epsilon) f_0'(\epsilon)\, dk_z, \\ &\text{where again } \epsilon = \epsilon(nk_z) = (n+1/2)\hbar\omega + (\hbar^2k_z^2/2m). \end{aligned}\right\} \tag{4.3}$$

Another useful expression obtains by writing the integrand as a function of the total energy of the electron ϵ, namely

$$\sigma = -\frac{2e^2}{m}\frac{1}{(2\pi)^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \times \hbar\omega \int_{(\hbar\omega/2)}^{\infty} \sum_n [\epsilon-(n+1/2)\hbar\omega]^{1/2}\tau(\epsilon) f_0'(\epsilon)\, d\epsilon, \tag{4.4}$$

where in the summation n goes over the non-negative integers that make $\epsilon-(n+1/2)\hbar\omega > 0$. Substituting the expression (3.7) for the relaxation time we get

$$\sigma(H) = -C \int_{(\hbar\omega/2)}^{\infty} \frac{\sum_n [\epsilon-(n+1/2)\hbar\omega]^{1/2}}{\sum_n [\epsilon-(n+1/2)\hbar\omega]^{-1/2}} f'_0(\epsilon)\, d\epsilon, \tag{4.5}$$

where

$$C = (\rho s^2/E_1^2 kT)(2e^2\hbar/\pi m).$$

For zero magnetic field we can calculate directly the conductivity. We find, using equation (3.9) for the relaxation time,

$$\sigma_0 = -C \int_0^{\infty} \frac{\epsilon}{3} f'_0(\epsilon)\, d\epsilon. \tag{4.6}$$

We can see that $\sigma(H)$, equation (4.5), reduces to σ_0, equation (4.6), in the limit of zero magnetic field. For, according to equation (3.10),

$$\lim_{H\to 0} \hbar\omega \sum_n [\epsilon-(n+1/2)\hbar\omega]^{-1/2} = 2\epsilon^{1/2}$$

and similarly

$$\lim_{H\to 0} \hbar\omega \sum_{n'} [\epsilon-(n'+1/2)\hbar\omega]^{1/2} = \int_0^{\epsilon} (\epsilon-x)^{1/2}\, dx = \tfrac{2}{3}\epsilon^{3/2}.$$

Therefore, the ratio of the two sums appearing in the integrand of equation (4.5) has the limiting value $(2/3)\epsilon^{3/2}/2\epsilon^{1/2} = \epsilon/3$.

The usual treatment of magnetoresistance which treats the influence of the magnetic field essentially as a classical Lorentz force yields for the case under consideration a conductivity independent of the magnetic field, and therefore equal to σ_0.

For finite magnetic fields the sums cannot be approximated by integrals and thus the quantization of the electron orbits in a magnetic field gives rise to a field dependent conductivity.

Since $\bar{v}_x = \bar{v}_y = 0$ we easily see that the usual magnetoresistance, as measured in a long wire with the magnetic and electric fields along the wire, is $\Delta\rho/\rho_0 = \sigma_0/\sigma(H) - 1$.

Below we shall discuss the conductivity $\sigma(H)$ in the limiting cases of degenerate and classical statistics.

A. Degenerate case

For this case the Fermi energy is much larger than kT. The fermi energy in zero magnetic field, ζ_0, is determined from the equation

$$n = 2\int_0^\infty N_0(\epsilon) f_0(\epsilon)\, d\epsilon = \frac{1}{3\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\zeta_0^{3/2}, \tag{4.7}$$

where n is the electron concentration. The conductivity is then

$$\sigma_0 = (e^2/m) n \tau_0(\zeta_0). \tag{4.8}$$

The Fermi energy in the presence of the magnetic field, ζ_H, is given through the equation

$$n = 2\int_{(\hbar\omega/2)}^\infty N(\epsilon) f_0(\epsilon)\, d\epsilon = -\int_{(\hbar\omega/2)}^\infty \left(\int_0^\epsilon 2N(\epsilon')\, d\epsilon'\right) f'_0(\epsilon) d\epsilon. \tag{4.9}$$

The distribution function $f_0(\epsilon)$ decreases from the value 1 to 0 within a range kT around the point $\epsilon = \zeta_H$. The density of states $N(\epsilon)$ now is not a smooth function, but has infinite discontinuities at the points $\epsilon = (n+1/2)\hbar\omega (n = 0, 1, 2, \ldots)$. If, therefore, in addition to $kT \ll \zeta_H$, we have $kT \ll \hbar\omega$, we can write

$$n = \frac{2}{(2\pi)^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \times \hbar\omega \sum_{n'=0}^{\bar{n}} [\zeta_H - (n'+1/2)\hbar\omega]^{1/2}, \tag{4.10}$$

where $\bar{n}$ is the largest integer for which

$$\zeta_H - (\bar{n}+1/2)\hbar\omega > 0.$$

The conductivity in these circumstances is

$$\sigma(H) = (e^2/m) n \tau(\zeta_H). \tag{4.11}$$

Thus the conductivity of a degenerate semiconductor exhibits, according to this simple model and for $\hbar\omega \gg kT$, an oscillatory variation with the magnetic field. As the Fermi level passes through an oscillator state the conductivity becomes zero. This magnetic field dependence reflects directly the influence of the magnetic field on the relaxation time. If the magnetic field had no effect on the relaxation time the conductivity would still show some field dependence due to the shifting of the fermi energy with the magnetic field. This, in general, is negligible, however.

For high magnetic fields, such that $(3/2)\hbar\omega > \zeta_H$, all electrons are in the ground oscillator state. This case was discussed in reference (1) in some detail. Equation (4.11) becomes in such case

$$\sigma(H) = \frac{4}{3}\sigma_0\left(\frac{\zeta_0}{\hbar\omega}\right)^2. \tag{4.12}$$

For arbitrary values of $\hbar\omega/kT$ but for $\hbar\omega/\zeta_H \ll 1$ we may use Poisson's summation formula to estimate the conductivity. The denominator in the integrand is given by equation (3.12) and the analogous expression for the numerator is found to be

$$\sum_n [\epsilon - (n+1/2)\hbar\omega]^{1/2} \cong \frac{1}{\hbar\omega}\frac{2}{3}\epsilon^{3/2}\left[1 - \frac{3}{4\pi\sqrt{2}}\left(\frac{\hbar\omega}{\epsilon}\right)^{3/2} \times \sum_1^\infty \frac{(-1)^r}{r^{3/2}} \cos\left(2\pi r\frac{\epsilon}{\hbar\omega} + \frac{\pi}{4}\right)\right]. \tag{4.13}$$

From equations (3.12) and (4.13) we note that for $\hbar\omega/\epsilon \ll 1$ the amplitude of oscillations of the relaxation time as a function of the magnetic field is larger than that of the other factors in the integrand. Thus the magnetic field dependence of the conductivity for small fields is determined through the field dependence of the relaxation time. For $\hbar\omega/\zeta_H \ll 1$ we therefore see that

$$\frac{\Sigma[\epsilon-(n+1/2)\hbar\omega]^{1/2}}{\Sigma[\epsilon-(n+1/2)\hbar\omega]^{-1/2}} \cong \frac{\epsilon}{3}\left[1-\left(\frac{\hbar\omega}{2\epsilon}\right)^{1/2}\sum_1^\infty \frac{(-1)^r}{r^{1/2}}\cos\left(2\pi r\frac{\epsilon}{\hbar\omega}-\frac{\pi}{4}\right)+\ldots\right]. \tag{4.14}$$

The integration of the fast varying terms $\cos[2\pi r(\epsilon/\hbar\omega)-(\pi/4)]$ can be carried out exactly with the accurate expression for $f_0'(\epsilon)$. For the smooth functions ϵ and $\epsilon^{1/2}$ we may take $f_0'(\epsilon) = -\delta(\epsilon-\zeta_H)$ We find

$$\sigma = \frac{C}{3}\zeta_H\left[1-\pi^2\sqrt{(2)}\frac{kT}{\hbar\omega}\left(\frac{\hbar\omega}{\zeta_H}\right)^{1/2}\sum_1^\infty (-1)^r r^{1/2}\frac{\cos\left(2\pi r\frac{\zeta_H}{\hbar\omega}-\frac{\pi}{4}\right)}{\sinh(2\pi^2 rkT/\hbar\omega)}\right]. \tag{4.15}$$

The Fermi energy ζ_H is obtained, in general, from the equation

$$n = \frac{(2m\zeta_H)^{3/2}}{3\pi^2\hbar^3}\left[1-\frac{3\pi\sqrt{2}}{4}\frac{kT}{\hbar\omega}\left(\frac{\hbar\omega}{\zeta_H}\right)^{3/2}\sum_1^\infty \frac{(-1)^r}{r^{1/2}}\frac{\cos\left(2\pi r\frac{\zeta_H}{\hbar\omega}+\frac{\pi}{4}\right)}{\sinh(2\pi^2 rkT/\hbar\omega)}\right]. \tag{4.16}$$

To the degrees of accuracy of equation (4.15), equation (4.16) gives

$$\zeta_H = \zeta_0 = \frac{(3\pi^2)^{2/3}\hbar^2}{2m}n^{2/3},$$

and therefore, with the use of equation (4.6),

$$\frac{\sigma(H)}{\sigma_0} \cong 1-\pi^2\sqrt{(2)}\frac{kT}{\hbar\omega}\times \left(\frac{\hbar\omega}{\zeta_0}\right)^{1/2}\sum_1^\infty (-1)^r r^{1/2}\frac{\cos\left(2\pi r\frac{\zeta_0}{\hbar\omega}-\frac{\pi}{4}\right)}{\sinh(2\pi^2 rkT/\hbar\omega)}. \tag{4.17}$$

For comparison the orbital magnetic susceptibility of the electron gas under the same conditions is[5]

$$\frac{\chi(H)}{\chi(0)} \cong 1-6\pi\sqrt{(2)}\frac{kT}{\hbar\omega}\times \left(\frac{\zeta_0}{\hbar\omega}\right)^{1/2}\sum_1^\infty \frac{(-1)^r}{r^{1/2}}\frac{\cos\left(2\pi r\frac{\zeta_0}{\hbar\omega}+\frac{\pi}{4}\right)}{\sinh(2\pi^2 rkT/\hbar\omega)},$$

where $\chi(0) = -n\mu_B^2/2\ \zeta_0$ is the steady susceptibility ($\mu_B = e\hbar/2mc$ = Bohr magneton). We note that the amplitude of the oscillations in the susceptibility is larger than that in the conductivity roughly by a factor $\zeta_0/\hbar\omega$.

B. Non-degenerate case

For classical statistics $f_0(\epsilon) = \exp[(\zeta-\epsilon)/kT]$ and $\zeta \ll kT$. The Fermi level in the presence of the magnetic field, ζ_H, as given by equation (4.9), is now obtained from

$$\exp(\zeta_H/kT) = n4\pi^{3/2}\times \left(\frac{\hbar^2}{2m}\right)^{3/2}(kT)^{-3/2}\frac{\sinh(\hbar\omega/2kT)}{(\hbar\omega/2kT)}. \tag{4.18}$$

The conductivity is then, according to equations (4.5) and (4.6),

$$\frac{\sigma(H)}{\sigma_0} = 3\frac{1-e^{-x}}{x}\int_0^\infty \frac{\sum_n (s-nx)^{1/2}}{\sum_n (s-nx)^{-1/2}}e^{-s}\,ds \tag{4.19}$$

with $x = \hbar\omega/kT$.

For $\hbar\omega/kT \gg 1$ most electrons are in the ground oscillator state and the "quantum limit" case obtains. As seen from equation (4.19), in this case

$$\sigma(H)/\sigma_0 = 3(kT/\hbar\omega). \tag{4.20}$$

This limiting case has been discussed in more detail in reference (1).

For $\hbar\omega/kT \ll 1$ the electrons occupy highly excited oscillator states. DR. LOWELL SCHOENFELD of these Laboratories was able to find, by use of the Euler summation formula with the remainder, that for $\hbar\omega/kT \ll 1$ equation (4.19) gives

$$\frac{\sigma(H)}{\sigma_0} = 1+17{\cdot}4\left(\frac{\hbar\omega}{kT}\right)^{1/2}+ \\ +\frac{3}{4}\frac{\hbar\omega}{kT}\ln\frac{\hbar\omega}{kT}+0(\hbar\omega/kT). \tag{4.21}$$

5. SUMMARY

The main result of this calculation is that for an electron in a magnetic field a relaxation time for collisions with acoustic lattice vibrations exists, which depends on the energy of the electron and is, in our approximation, inversely proportional to the density of states. Thus the effect of the magnetic field on the relaxation time is given directly through its effect on the density of states. As a function of energy the relaxation time presents an oscillatory behavior with "period" $\hbar\omega$. These oscillations can be seen in the conductivity only if the electron gas is degenerate. The amplitude of these oscillations, however, is much smaller than the corresponding amplitude of the de Haas–van Alphen oscillations in the magnetic susceptibility. For non-degenerate statistics the oscillations disappear. An over-all variation of the conductivity with the magnetic field remains in any case and this represents a quantum mechanical effect, since the classical theory predicts zero magnetoresistance for such an isotropic electron gas.

As has been mentioned,[1] a substance that seems suitable for investigating these points is InSb, or any other semiconductor or semimetal with small effective mass and simple energy surface. In view of the restrictions on temperature for the observability of the oscillations and the fact that these formulae are valid only in lattice scattering range a highly pure specimen is desirable. However, it is not impossible that impurity scattering too, might give a relaxation time with oscillatory dependence on the magnetic field.

VOL.
4
1958

Acknowledgements—It is a pleasure to express my gratitude to Dr. LOWELL SCHOENFELD for the evaluation of the integral appearing in equation (4.19) for small x, and to Dr. E. N. ADAMS for his comments on the manuscript.

REFERENCES

1. ARGYRES P. N. and ADAMS E. N. *Phys. Rev.* **104,** 900 (1956).
2. TITEICA V. S. *Ann. Phys., Paris* **22,** 129 (1935).
3. DAVYDOV B. and POMERANCHUK I. *J. Phys., Moscow* **2,** 147 (1940).
4. COURANT R. and HILBERT D. *Methods of Mathematical Physics* Vol. 1 p. 76. Interscience, New York (1953).
5. SONDHEIMER E. H. and WILSON A. H. *Proc. Roy. Soc.* A **210,** 173 (1951).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 27–49.

WAHRSCHEINLICHKEITSVERTEILUNG UND KORRELATION VON SCHUBSPANNUNGEN IM KRISTALL

GERHARD BARSCH

Institut für theoretische Physik der Universität Göttingen*

(*Received* 1 *April* 1957)

VOL. 4 958

Zusammenfassung—Für die lokalen statistischen Variablen im Kristall (Verschiebungen, Verzerrungen, Spannungen) werden die Wahrscheinlichkeitsverteilungen angegeben. Die darin auftretenden Schwankungsquadrate und die räumlichen Korrelationsfunktionen werden mittels Debyescher Näherung für das kubisch flächenzentrierte Gitter berechnet. Die Reichweite der Korrelation ist von der Grössenordnung einer Gitterkonstanten.

Sodann wird das Verhalten der Volumenmittelwerte der lokalen statistischen Variablen und insbesondere der Einfluss der Gestalt des Probevolumens und der Korrelation am Beispiel der Schubspannungen untersucht. Es zeigt sich, dass das Schwankungsquadrat der Volumenmittelwerte im klassischen Grenzfall hoher Temperatur umgekehrt proportional vom Probevolumen abhängt, während dessen Gestalt und die Korrelation keinen Einfluss auf das Resultat haben. Die am absoluten Nullpunkt vorhandenen Nullpunktsschwankungen verhalten sich für Probevolumina, die in mindestens einer Dimension nur atomare Ausdehnung besitzen, in bezug auf Volumen- und Korrelationsabhängigkeit wie die klassischen thermischen Schwankungen. Für makroskopische Volumina ist jedoch die Volumenabhängigkeit stärker und der Einfluss der Korrelation negativ. Diese Ergebnisse werden durch die Berücksichtigung des gittertheoretischen Spektrums nicht wesentlich geändert.

Als Anwendung wird eine statistische Begründung der Beckerschen Formel, welche die Wahrscheinlichkeitsverteilung für die Schubspannung in einem Probevolumen angibt, für beliebige Probevolumina gebracht. Die strenge statistische Begründung bestätigt für hohe Temperaturen im wesentlichen die Beckersche Formel, liefert aber im Gegensatz dazu am absoluten Nullpunkt ein von der Nullpunktsbewegung herrührendes Schwankungsquadrat, das für Volumina von atomarer Ausdehnung mit dem Beitrag der klassischen Wärmebewegung vergleichbar ist.

Abstract—The distribution functions for the local statistical variables in a crystal (displacements, strains, stresses) are given and the second moments calculated by means of Debye's approximation. The range of correlation is about one lattice constant.

The variances of new variables obtained by averaging the local variables over some fixed volume are calculated for the case of shearing stresses. In the classical limit these variances increase proportionally to the reciprocal volume and are independent of the shape of the volume and of the correlation of the local variables. At absolute zero the zero-point fluctuations behave similarly if the extension of the volume in one or more dimensions does not exceed a few lattice constants. For volumes of macroscopic size the variances increase more rapidly than proportionally with reciprocal volume and the influence of the correlation becomes negative. These results essentially remain unaltered if the lattice spectrum is considered.

As an application, Becker's formula, denoting the distribution function for the shearing stresses in a volume, is derived from statistical principles. The formula obtained in the classical limit essentially agrees with the common thermodynamic expression of Becker's formula. For absolute zero, however, it leads to an expression of the variance of the zero-point motion which, for volumes of atomic size, is comparable with the contribution of classical thermal motion.

* Jetzige Anschrift : Battelle Institut, Frankfurt/Main.

1. EINLEITUNG: DIE BECKERSCHE FORMEL

(a) *Thermodynamische Ableitung*

EINER der ersten Ansätze zu einer Theorie der plastischen Verformung von Metallkristallen wurde von BECKER gegeben.[1] BECKER ging dabei von der Vorstellung aus, dass das plastische Fliessen eines Kristalls unter dem Einfluss einer äusseren Schubspannung durch die Wärmebewegung des Kristalls begünstigt wird. Danach sollen etwa zwei Netzebenen immer dann gegeneinander abgleiten, wenn in einem gewissen Volumen v die äussere und die durch thermische Schwankungen entstandene Schubspannung zusammen den für die Abgleitung erforderlichen kritischen Wert überschreiten. Die Wahrscheinlichkeit für dieses Ereignis lässt sich sofort aus der „Beckerschen Formel" berechnen. Die Beckersche Formel gibt die Wahrscheinlichkeit dafür an, dass die durch thermische Schwankungen entstandene Schubspannung eines Probevolumens v in dem Intervall $\tau,\ \tau+d\tau$ liegt und lautet:

$$W(\tau)\,d\tau = \frac{1}{\sqrt{(2\pi\overline{\tau^2})}}\exp\left(-\frac{\tau^2}{2\overline{\tau^2}}\right)d\tau. \tag{1.1a}$$

Die Wahrscheinlichkeitsverteilung ist auf eins normiert:

$$\int_{-\infty}^{+\infty} W(\tau)\,d\tau = 1.$$

Das Schwankungsquadrat beträgt*

$$\overline{\tau^2} = g\frac{kT}{v}. \tag{1.1b}$$

Hierbei bedeutet g den Schubmodul des als isotrop angenommenen Mediums, k die Boltzmannsche Konstante und T die absolute Temperatur.

Zur Begründung der Beckerschen Formel betrachte man einen Kristall vom Volumen V und darin ein quaderförmiges Teilvolumen v, das klein gegenüber dem Gesamtvolumen V sein soll. Diese ganze Anordnung befinde sich in einem Wärmebad der Temperatur T. Unter dem Einfluss der thermischen Bewegung ist die Gestalt des Probevolumens v, die durch die homogene Scherung $\varepsilon = \varepsilon_{12}$ gekennzeichnet sein möge, Schwankungen unterworfen. Dabei soll $\varepsilon_{12} = \frac{1}{2}(\partial s_1/\partial x_2 + \partial s_2/\partial x_1)$ sein, wobei im Hinblick auf die Erzeugung oder Bewegung von Versetzungen x_1 in die Gleitrichtung, x_2 in die Richtung der Gleitebenennormalen des betreffenden Kristalls fallen und s_i die Komponenten des Verschiebungsvektors in der x_i-Richtung bezeichnen möge. Fasst man ε als Parameter auf, der makroskopisch fixiert werden kann, so lässt sich die freie Energie des ganzen Kristalls in die Beiträge der beiden Teilvolumina v und $V-v$ aufspalten:

$$F(T,V,\epsilon) = F_1(T,v,\epsilon)+F_2(T,V-v,[\epsilon(\boldsymbol{\nu})]).$$

Eigentlich tritt auf der rechten Seite auch noch ein Wechselwirkungsglied F_{12} hinzu, das jedoch vernachlässigt werden kann.

F_2 ist ein Funktional des Verzerrungsfeldes $\varepsilon(\boldsymbol{\nu})$ in dem Volumen $V-v$. F_1 und F_2 lassen sich wiederum in die gewöhnliche freie Energie und die von der elastischen Deformation herrührenden Beiträge zerlegen:

$$F_1(T,v,\epsilon) = F_1(T,v,0)+2g\epsilon^2 v$$

$$F_2(T,V-v,[\epsilon(\boldsymbol{\nu})]) = F_2(T,V-v,0)+2g\int_{V-v}\epsilon^2(\boldsymbol{\nu})\,d\boldsymbol{\nu}$$

Für $V \gg v$ ist die elastische Energie $2g\int\epsilon^2(\boldsymbol{\nu})\,d\boldsymbol{\nu}$ ausserhalb des homogen verzerrten Probevolumens v klein gegenüber dessen Beitrag $2g\epsilon^2 V$, sodass man für die gesamte freie Energie einfach

$$F(T,V,\epsilon) = F_1(T,v,0)+F_2(T,V-v,0)+2g\epsilon^2 v$$

erhält.

Die Wahrscheinlichkeit, im Probevolumen v eine Verzerrung in dem Intervall $\varepsilon,\varepsilon+d\varepsilon$ vorzufinden, beträgt

$$W(\epsilon)\,d\epsilon = C\exp[-F(T,V,\epsilon)/kT]\,d\epsilon$$
$$= C'\exp[-(2gV/kT)\epsilon^2]\,d\epsilon.$$

Hierbei sind C und C' Normierungskonstante.

Die Scherung des Probevolumens ist mit einer elastischen Spannung $\tau = 2g\epsilon$ verknüpft, für deren Wahrscheinlichkeitsverteilung man hieraus durch Einsetzen sofort Gleichung (1.1) erhält.†

Berechnet man aus der Formel (1.1) die Fliessgeschwindigkeit und die elastische Grenze,

* Der Strich bedeutet hier und im folgenden stets die Mittelung über die zugehörige Verteilung, also

$$\overline{\tau^2} = \int_{-\infty}^{+\infty}\tau^2 w(\tau)\,d\tau.$$

† Sofern sämtliche äusseren Spannungen an dem Kristall null sind. Steht der Kristall jedoch unter dem Einfluss äusserer (Oberflächen- oder Volumen-) Kräfte, so erhält man die Verteilung der Spannung aus einer zu der obigen analogen Formel, wo statt der freien Energie die freie Enthalpie auftritt. Wir setzen im folgenden immer Abwesenheit äussere Kräfte voraus, sodass freie Energie und freie Enthalpie gleich werden. Dann sind die Verteilungen für die Verzerrungen und für die Spannungen identisch.

VOL. 4 1958

so ergibt sich beim Vergleich mit dem Experiment für das Probevolumen v ein viel zu kleiner Wert, der auch vom Standpunkt der Erzeugung von Versetzungen ganz unverständlich ist.[2] Es war nun zu vermuten, dass diese Schwierigkeit durch eine ungerechtfertigte Anwendung der Formel (1.1) entsteht. Einerseits beruht nämlich die Beckersche Formel auf der Thermodynamik und gilt daher nur für grosse Probevolumina. Andererseits ist für die thermisch aktivierte Abgleitung zweier Netzebenen nur erforderlich, dass die Schubspannung in einem sehr flachen, in der Gleitebene gelegenen Probevolumen den kritischen Wert überschreitet. Auf solche flachen Probevolumina, deren Höhe etwa nur eine Gitterkonstante beträgt, lässt sich aber die thermodynamisch begründete Formel (1.1) nicht anwenden. Man gelangt somit zu der Aufgabe, anstelle von (1.1) eine für beliebige Volumina gültige Verteilung mit statistischen Mitteln abzuleiten.

(b) *Statistische Begründung*

Jede statistische Begründung muss von der Verteilung $w(\tau_1, ..., \tau_n)$ der lokalen Schubspannungen im Kristallgitter ausgehen, wobei $w(\tau_1, ..., \tau_n)d\tau_1 ... d\tau_n$ die Wahrscheinlichkeit dafür angibt, dass in den Gitterpunkten $\boldsymbol{\nu}_1, ..., \boldsymbol{\nu}_n$ die thermische Schubspannung in dem Intervall $\tau_1, ..., \tau_n$; $\tau_1+d\tau_1, ..., \tau_n+d\tau_n$ liegt.* Die Atome sollen in dem Probevolumen v liegen. $v_0 = v/n$ ist das Volumen pro Atom und beträgt beim kubisch flächenenzentrierten Gitter, das im folgenden allein betrachtet werden soll, $v_0 = a^3/4$. Dabei ist a die Kantenlänge des Elementarwürfels.

Wie bei der oben angeführten thermodynamischen Ableitung der Beckerschen Formel wird man zunächst nach der Wahrscheinlichkeit dafür fragen, dass das Probevolumen v eine bestimmte, durch die Verzerrung ϵ gekennzeichnete Gestalt besitzt. Bezeichnet $w(\epsilon_1, ..., \epsilon_n)$ die Verteilungsfunktion der Verzerrungen in den n Gitterpunkten des Probevolumens, so liegt es nahe, zur Beschreibung der Gestalt des ganzen Probevolumens die Wahrscheinlichkeit

$$\int_{\epsilon \leqq \epsilon_i \leqq \epsilon+d\epsilon} \cdots \int W(\epsilon_1, ..., \epsilon_n)\, d\epsilon_1 \ldots d\epsilon_n = W(\epsilon, ..., \epsilon)\, d\epsilon^n$$

zu verwenden. Dieser Ausdruck gibt die Wahrscheinlichkeit dafür an, dass in allen n Gitterpunkten die Verzerrungen in dem Intervall ϵ, $\epsilon+d\epsilon$ liegen. Da diese Wahrscheinlichkeit in dem Differential $d\epsilon$ vom n ten Grade ist, lässt sie sich nicht mit der der Beckerschen Formel zugrundeliegenden Verteilung $w(\epsilon)\, d\epsilon$ vergleichen. Die gleiche Schwierigkeit liegt auch dann vor, wenn man diese Betrachtung nur auf die Oberflächenatome des Quaders anwendet. Der entscheidende Schritt besteht nun darin, dass die über das Probevolumen gemittelte Verzerrung

$$\overline{\epsilon_{12}}^{v} = \frac{1}{n} \sum_{i=1}^{n} \epsilon_{12}{}^{i}$$

betrachtet wird. Hierbei sollen $\epsilon_{12}{}^{i}$ die Verzerrungen in demjenigen Verschiebungsfeld $\boldsymbol{s}(\boldsymbol{\nu})$ sein, das die Verschiebungen im Gitter stetig interpoliert, und zwar genommen an den Gitterpunkten.† Die Summe ist über die n Atome des Probevolumens zu erstrecken, das im folgenden stets quaderförmig angenommen werden soll. Die mittlere Verzerrung $\epsilon = \overline{\epsilon_{12}}^{v}$ ist gerade der bei der obigen Ableitung der Beckerschen Formel als fixierbar vorausgesetzte Parameter.

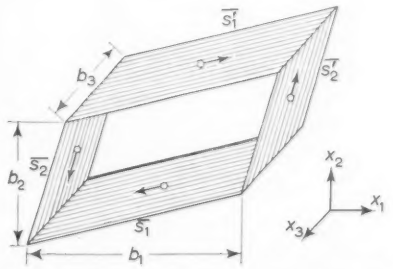


ABB. 1. Zur Beschreibung der Verzerrung eines Quaders mit den Kantenlängen b_1, b_2, b_3. Die Pfeile bezeichnen die Verschiebungskomponenten der Schwerpunkte der jeweiligen Flächen.

* Zunächst sind ja elastische Spannungen nur in einem elastischen Kontinuum erklärt. Indessen lassen sich auch im Gitter Spannungen definieren, wenn man die Verschiebungen der Atome durch ein stetiges Verschiebungsfeld interpoliert und dieses fiktive Verschiebungsfeld anstelle des elastischen differenziert. Man erhält so ein stetiges Spannungsfeld, dessen Werte aber nur an den Gitterpunkten interessieren. Diese Werte sind in $w(\tau_1, ..., \tau_n)$ gemeint.

† Wie oben beziehen sich die Indizes auf eine Basis, deren x_1-bezw. x_2-Achse in die Gleitrichtung bezw. Gleitebenennormale des betreffenden Kristalls fällt.

Die Bedeutung der mittleren Verzerrung ϵ lässt sich einsehen, wenn die Summe durch ein Integral über das Volumen $v = b_1 b_2 b_3$ des Quaders (Abb. 1) ersetzt und das Integral nach dem Gausschen Satz umgeformt wird. Man findet so, dass ϵ das Mittel der Verschiebungsdifferenzen der beiden gegenüberliegenden Schwerpunktspaare bedeutet :

$$\epsilon \equiv \overline{\epsilon_{12}}^{v} = \frac{1}{2}\left(\frac{\overline{s_1'} - \overline{s_1}}{b_2} + \frac{\overline{s_2'} - \overline{s_2}}{b_1}\right).$$

Hierbei sind $\overline{s_1'}$ und $\overline{s_1}$ bezw. $\overline{s_2'}$ und $\overline{s_2}$ die Verschiebungen der Schwerpunkte der im Abstand b_2 bezw. b_1 gegenüberliegenden Grenzflächen (Abb. 1). Dies gilt zwar nur dann, wenn die Gittersumme durch ein Integral ersetzt werden darf, also für grosse Probevolumina; aber offenbar ist auch für flache Quader, bei denen diese anschauliche Deutung nicht möglich ist, die Verwendung von ϵ eine sehr plausible Möglichkeit zur Beschreibung der Gestalt des Probevolumens.

Tatsächlich interessiert jedoch in erster Linie nicht die Gestalt, sondern die in dem Probevolumen wirksamen Spannungen. Diese sind der Verzerrung proportional, und es ist deshalb wieder—bei Abwesenheit äusserer Spannungen—genau dasselbe, ob man die Verzerrungen oder die dazu proportionalen Spannungen verwendet. Insbesondere lässt sich das Probevolumen als Ganzes auch durch die mittlere Schubspannung $\overline{\tau_{12}}^{v} = g\overline{\epsilon_{12}}^{v}$, also durch

$$\tau \equiv \overline{\tau_{12}}^{v} = \frac{1}{n}\sum_{i=1}^{n} \tau_i \qquad (1.2)$$

beschreiben.

Die in der Beckerschen Formel auftretende Spannung hat also gerade die Bedeutung der mittleren Schubspannung des Probevolumens. Demnach ergibt sich die statistisch begründete Formel hierfür zu

$$w(\tau) = \int_{-\infty}^{+\infty} \cdots \int \delta\left(\tau - \frac{1}{n}\sum_{i=1}^{n} \tau_i\right) w(\tau_1, \ldots, \tau_n)\, d\tau_1 \ldots, d\tau_n. \qquad (1.3)$$

Hierbei bezeichnet $\delta(x)$ die Diracsche δ-Funktion.

Die noch verbleibende Aufgabe besteht nun in der Ableitung der Funktion $w(\tau_1, \ldots, \tau_n)$. Darin wie in (1.3) geht die statistische Abhängigkeit der einzelnen Variablen explizit ein, und man kann nun die Frage nach der Grösse dieser Abhängigkeit sowie nach deren Einfluss auf die Verteilung (1.3) aufwerfen. Man wird also bei der Begründung der Beckerschen Formel dazu geführt, die Verteilungsfunktion $w(\tau_1, \ldots, \tau_n)$ abzuleiten und die Natur der darin auftretenden Korrelation zu untersuchen. Da die Spannungen im wesentlichen die Differentialquotienten der Verschiebungen sind, liegt es nahe, die Untersuchung auch auf die Verschiebungen selbst auszudehnen und nach der Wahrscheinlichkeitsverteilung $w(\ldots \boldsymbol{s_1}^m, \boldsymbol{s_2}^m, \boldsymbol{s_3}^m, \ldots) \ldots ds_1^m ds_2^m ds_3^m \ldots$ für die Verschiebungskomponenten $\ldots s_1^m, s_2^m, s_3^m \ldots$ zu fragen. Das soll im folgenden Abschnitt geschehen, obwohl nur ein Teil der darin auftretenden Ergebnisse für die Ableitung und Diskussion der Formel (1.3) benötigt wird.

2. DIE LOKALEN STATISTISCHEN VARIABLEN IM KRISTALL

(a) *Normalkoordinaten**

Wir beschränken uns von vornherein auf das kubisch flächenzentrierte Gitter, obgleich die meisten Formeln allgemein gelten werden. Wir gehen aus von einem Periodizitätsvolumen im Kristallgitter, das N Atome der Masse m enthalten möge. Der mechanische Zustand des Kristalls wird dann durch Angabe sämtlicher Atomlagen $\boldsymbol{R}^m$ in Abhängigkeit von der Zeit beschrieben. Wir verwenden stattdessen die Verschiebungen $\boldsymbol{s}^m$ aus den Ruhelagen $\boldsymbol{\nu}^m$, die durch $\boldsymbol{s}^m = \boldsymbol{R}^m - \boldsymbol{\nu}^m$ gegeben sind. Die Ruhelagen sind durch $\boldsymbol{\nu}^m = (A)\boldsymbol{m}$ gegeben, wobei $\boldsymbol{m}$ ein Vektor mit ganzzahligen Komponenten und (A) die aus den Komponenten der Elementarzelle $\boldsymbol{a}_\kappa$ gebildete Matrix ist, nämlich

$$(A) = \frac{a}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \text{ (Abb. 2).}$$

Die $3N$ Komponenten s_i^m der Verschiebungsvektoren $\boldsymbol{s}^m = (s_1^m, s_2^m, s_3^m)$ lassen sich in eine Fourier-Reihe nach den $3N$ „Normalkoordinaten" $q_{\boldsymbol{k}\lambda}$ entwickeln :

$$s_i^m = \sum_{\boldsymbol{k}\lambda} e_{\lambda i}^{\boldsymbol{k}}\, \alpha_m^{\boldsymbol{k}}\, q_{\boldsymbol{k}\lambda}. \qquad (2.1)$$

Die Summe über $\boldsymbol{k}$ ist über die N Werte des Wellenzahlvektors der 1. Brillouin-Zone zu erstrecken, λ läuft über die 3 Polarisationsrichtungen.

* Für die gittertheoretische Grundlage siehe z.B. LEIBFRIED.[3]

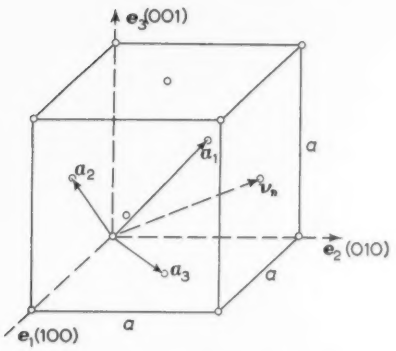


ABB. 2. Die Lage der primitiven Basis e_1, e_2, e_3 im kubisch flächenzentrierten Gitter.

Allgemein gilt $\boldsymbol{v}^{\boldsymbol{m}} = (A)\boldsymbol{m}$, wobei (A)

$$(A) = \frac{a}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Für den eingezeichneten Gittervektor $\boldsymbol{v}^{\boldsymbol{m}}$ ist $m = (1, 0, 1)$

Die Komponenten der 3 Polarisationsvektoren $\boldsymbol{e}_\lambda{}^{\boldsymbol{k}} = \{e_{\lambda 1}{}^{\boldsymbol{k}}, e_{\lambda 2}{}^{\boldsymbol{k}}, e_{\lambda 3}{}^{\boldsymbol{k}}\}$ bilden eine orthogonale Matrix :

$$\sum_{\lambda=1}^{3} e_{\lambda i}{}^{\boldsymbol{k}} e_{\lambda j}{}^{\boldsymbol{k}} = \delta_{ij}$$

$$\sum_{i=1}^{3} e_{\lambda i}{}^{\boldsymbol{k}} e_{\mu i}{}^{\boldsymbol{k}} = \delta_{\lambda\mu}. \qquad (2.2a)$$

Auch die $\alpha_{\boldsymbol{m}}{}^{\boldsymbol{k}}$ bilden ein normiertes Orthogonalsystem. Es ist

$$\alpha_{\boldsymbol{m}}{}^{\boldsymbol{k}} = \begin{cases} \sqrt{\left(\frac{2}{N}\right)} \cos \boldsymbol{k}\boldsymbol{v}_{\boldsymbol{m}} & k_1 > 0 \\ \sqrt{\left(\frac{2}{N}\right)} \sin \boldsymbol{k}\boldsymbol{v}_{\boldsymbol{m}} & k_1 < 0 \\ \frac{1}{\sqrt{N}} & \boldsymbol{k} = 0 \end{cases} \qquad (2.2)$$

und

$$\sum_{k} \alpha_{\boldsymbol{m}}{}^{\boldsymbol{k}} \alpha_{\boldsymbol{n}}{}^{\boldsymbol{k}} = \delta_{\boldsymbol{m}\boldsymbol{n}}$$

$$\sum_{\boldsymbol{n}} \alpha_{\boldsymbol{n}}{}^{\boldsymbol{k}} \alpha_{\boldsymbol{n}}{}^{\boldsymbol{k}'} = \delta_{\boldsymbol{k}\boldsymbol{k}'} \qquad (2.2c)$$

(2.1) stellt also insgesamt eine orthogonale Transformation dar. Sie separiert die Hamilton-Funktion des Kristalls in eine Summe von harmonischen Oszillatoren. Die kanonische Verteilung der Normalkoordinaten ist durch die Slater-Summe für den linearen Oszillator gegeben. Sie wurde von Husimi berechnet und liefert für $\boldsymbol{k} \neq 0$ die Verteilung[4]

$$w(q_{\boldsymbol{k}\lambda})\, dq_{\boldsymbol{k}\lambda} = \frac{1}{\sqrt{2\pi}\ \sigma_{\boldsymbol{k}\lambda}} \exp\left(-\frac{q_{\boldsymbol{k}\lambda}{}^2}{2\sigma_{\boldsymbol{k}\lambda}{}^2}\right) dq_{\boldsymbol{k}\lambda} \qquad (2.3a)$$

mit dem Schwankungsquadrat

$$\sigma_{\boldsymbol{k}\lambda}{}^2 = \overline{q_{\boldsymbol{k}\lambda}{}^2} = \frac{\epsilon_{\boldsymbol{k}\lambda}}{m\omega_{\boldsymbol{k}\lambda}{}^2} \qquad (2.3b)$$

$\epsilon_{\boldsymbol{k}\lambda}$ ist die mittlere thermische Energie eines Oszillators der Frequenz $\omega_{\boldsymbol{k}\lambda}$:

$$\epsilon_{\boldsymbol{k}\lambda} = \frac{\hbar\omega_{\boldsymbol{k}\lambda}}{2} \operatorname{ctgh} \frac{\hbar\omega_{\boldsymbol{k}\lambda}}{2kT} \qquad (2.4)$$

$\omega_{\boldsymbol{k}\lambda}$ sind die aus der Säkulargleichung des Kristalls zu bestimmenden Eigenfrequenzen. Wir werden uns bei den folgenden numerischen Beispielen auf Debyesche Näherung* beschränken und zudem Isotropie voraussetzen. Dann sind die Eigenfrequenzen gegeben durch

$$\omega_{\boldsymbol{k}\lambda} = c_\lambda k. \qquad (2.5)$$

Dabei ist $k = |\boldsymbol{k}|$. c_λ sind die Schallgeschwindigkeiten der Longitudinal- und Transversalwellen ($c_1 = c_{\text{long}}$, $c_2 = c_3 = c_{\text{tr}}$).

Der Wert $k = 0$ liefert die Schwerpunktsbewegung des Periodizitätsvolumens und ist gesondert zu betrachten.

Wegen (2.1) und (2.2) gilt

$$q_{0\lambda} = \Big(\sum_{i=1}^{3} e_{\lambda i}{}^{\boldsymbol{k}} \sum_{\boldsymbol{n}} s_i{}^{\boldsymbol{n}}\Big)_{k=0}.$$

Wird der Schwerpunkt festgehalten, was im folgenden stets vorausgesetzt werden soll, so ist $(\sum_{\boldsymbol{n}} s_i{}^{\boldsymbol{n}}) = 0$ zu setzen. Dann ist die Verteilung der $q_{0\lambda}$ jeweils eine δ-Funktion :

$$w(q_{0\lambda})\, dq_{0\lambda} = \delta(q_{0\lambda})\, dq_{0\lambda}. \quad (\lambda = 1, 2, 3)$$

Verwendet man für die δ-Funktion die Darstellung

$$\delta(x) = \lim_{\beta\to 0} \frac{1}{\sqrt{2\pi}\beta} \exp\left(-\frac{x^2}{2\beta^2}\right),$$

* Hier und im folgenden ist stets die erweiterte Debyesche Näherung gemeint, bei der für die Longitudinal- und Transversalwellen verschiedene Abschneidefrequenzen verwendet werden.

so bleibt formal (2.3a) auch für $\boldsymbol{k} = 0$ richtig, wo nur anstelle von (2.3b)

$$\sigma_{0\lambda}{}^2 = 0. \tag{2.3c}$$

gilt. Dann ist also der Fall $\boldsymbol{k} = 0$ formal auch in (2.3a) enthalten.

Die $3N$ Normalkoordinaten sind statistisch unabhängig, sodass man als Verteilungsfunktion für alle Koordinaten $q_{k\lambda}$ das Produkt über die Verteilungen (2.3a) erhält :

$$w(\dots q_{k\lambda}\dots)\dots dq_{k\lambda}\dots = \prod_{k\lambda}\left\{\frac{1}{\sqrt{2\pi}\,\sigma_{k\lambda}}\exp\left(-\frac{q_{k\lambda}{}^2}{2\sigma_{k\lambda}{}^2}\right)dq_{k\lambda}\right\}. \tag{2.6a}$$

Wegen $\omega_{k\lambda} = \omega_{-k\lambda}$ ist $\sigma_{k\lambda}{}^2 = \sigma_{-k\lambda}{}^2$, sodass man stattdessen auch

$$w(\dots q_{k\lambda}\dots)\dots dq_{k\lambda}\dots = \prod_{k_1>0}\left\{\frac{1}{2\pi\sigma_{k\lambda}{}^2}\times \exp\left[-\frac{1}{2\sigma_{k\lambda}{}^2}(q_{k\lambda}{}^2+q_{-k\lambda}{}^2)\right]dq_{k\lambda}dq_{-k\lambda}\right\} \tag{2.6b}$$

schreiben kann; das Produkt darf aber dann nur über die halbe Brillouin-Zone erstreckt werden.

Die Verteilungsfunktion (2.6) für die Normalkoordinaten beschreibt das gesamte statistische Verhalten des Kristalls. Für jede andere von den Normalkoordinaten abhängige Grösse lässt sich daraus die Wahrscheinlichkeitsverteilung ableiten.

Die Verteilung (2.6) hat den Vorteil, dass die darin auftretenden Variablen $q_{k\lambda}$ statistisch unabhängig sind. Ihr Nachteil liegt darin, dass die $q_{k\lambda}$ als Amplituden von Schallwellen räumlich nicht lokalisiert sind, sondern sich auf das ganze Periodizitätsvolumen des Kristalls beziehen. Die wirklichen geometrischen Schwankungserscheinungen im Kristall können aber nur durch die Verschiebungen der Atome oder davon abgeleitete Grössen beschrieben werden.

(b) *Verschiebungen*

Das statistische Verhalten des Kristalls lässt sich also auch durch Angabe der Wahrscheinlichkeitsverteilung für die Verschiebungen $s_i m$ beschreiben. Da die $s_i m$ nach (2.1) von den $q_{k\lambda}$ durch eine orthogonale Transformation abhängen, ergibt sich bei festgehaltenem Schwerpunkt des Periodizitätsvolumens aus der Verteilung (2.6) der $q_{k\lambda}$ durch Einsetzen von (2.1) die Verteilungsfunktion

$$w(\dots s_i m\dots)\dots ds_i m\dots = \lim_{\sigma_{0\lambda}\to 0}\frac{1}{(2\pi)^{3N/2}|S|^{1/2}}\times \exp\left(-\tfrac{1}{2}\sum_{\substack{mn\\ i\,k}} A_{ik}{}^{mn}s_i m s_k n\right)\dots ds_i m\dots \tag{2.7}$$

Die im Exponenten auftretende Matrix (A) ist gegeben durch

$$A_{ik}{}^{mn} = \sum_{k\lambda}\frac{1}{\sigma_{k\lambda}{}^2}e_{\lambda i}{}^k e_{\lambda k}{}^k \alpha_m{}^k \alpha_n{}^k. \tag{2.8a}$$

Die $s_i m$ sind statistisch abhängig. Ein Mass hierfür ist die „Korrelationsmatrix" (S), für deren Elemente man aus (2.1) wegen der statistischen Unabhängigkeit der $q_{k\lambda}$ sofort erhält :

$$S_{ik}{}^{mn} = \overline{s_i m s_k n} = \sum_{k\lambda}\sigma_{k\lambda}{}^2 e_{\lambda i}{}^k e_{\lambda k}{}^k \alpha_m{}^k \alpha_n{}^k \tag{2.8b}$$

$|S|$ ist die Determinante der Korrelationsmatrix. Da die Determinante der Transformation (5) eins ist, ergibt sich dafür

$$|S| = \mathrm{Det}(S) = \prod_{k\lambda}\sigma_{k\lambda}{}^2. \tag{2.8c}$$

Weiter gilt noch

$$(S) = (A^{-1}).$$

Zieht man aus den Summen und dem Produkt den Wert $\boldsymbol{k} = 0$ einzeln heraus und führt dann den Grenzübergang $\sigma_{0\lambda}\to 0$ aus, so erhält man bei Beachtung von (2.1) und der Exponentialdarstellung der δ-Funktion

$$w(\dots s_i m\dots)\dots ds_i m\dots = \frac{1}{(2\pi)^{(3N-3)/2}|S'|^{1/2}}\times \exp\left(-\tfrac{1}{2}\sum_{\substack{mn\\ i\,k}} A'_{ik}{}^{mn}s_i m s_k n\right)\prod_{j=1}^{3}\delta\left(\sum_n s_j n\right)\dots ds_i m\dots \tag{2.9}$$

Die Matrix (A') und die Determinante S' sind durch (2.8a) bzw. (2.8c) gegeben, wo nur jetzt der

Wert $\boldsymbol{k} = 0$ bei der Summe bzw. beim Produkt auszulassen ist. Die Korrelationsmatrix (S) bleibt wegen (2.3c) ungeändert, ist aber nicht die Reziproke der im Exponenten stehenden Matrix (A').

Einsetzen der Werte aus (2.2b) und (2.3b) in (2.8b) ergibt, wenn noch $\sigma_{\boldsymbol{k}\lambda}{}^2 = \sigma_{-\boldsymbol{k}\lambda}{}^2$, $e_{\lambda i}{}^{\boldsymbol{k}} e_{\lambda k}{}^{\boldsymbol{k}} = e_{\lambda i}{}^{-\boldsymbol{k}} e_{\lambda k}{}^{-\boldsymbol{k}}$ und $\alpha_m{}^{\boldsymbol{k}} \alpha_n{}^{\boldsymbol{k}} + \alpha_m{}^{-\boldsymbol{k}} \alpha_n{}^{-\boldsymbol{k}} = (2/N) \cos \boldsymbol{k}\boldsymbol{\nu}_{m-n}$ beachtet und die Summe dann nur über die halbe Brillouin-Zone erstreckt wird,

$$S_{ik}{}^{mn} = \frac{2}{N} \sum_{k_\iota > 0} s_{ik}{}^{\boldsymbol{k}} \cos \boldsymbol{k}\boldsymbol{\nu}_{m-n} \qquad (2.10)$$

Dabei bedeutet

$$s_{ik}{}^{\boldsymbol{k}} = \sum_{\lambda} \frac{\epsilon_{\boldsymbol{k}\lambda}}{m\omega_{\boldsymbol{k}\lambda}{}^2} e_{\lambda i}{}^{\boldsymbol{k}'} e_{\lambda k}{}^{\boldsymbol{k}}. \qquad (2.11)$$

Ein entsprechender Ausdruck ergibt sich für $A_{ik}{}^{mn}$. Die Diagonalelemente $S_{ii}{}^{mm}$ sind von $\boldsymbol{m}$ unabhängig. Bei kubischen Gittern gilt überdies aus Symmetriegründen $S_{11}{}^{mm} = S_{22}{}^{mm} = S_{33}{}^{mm}$, sodass hier die Diagonalelemente der Korrelationsmatrix alle den gleichen Wert besitzen.

Die Diagonalelemente stellen die Schwankungsquadrate der Verschiebungskomponenten, bezogen auf den Schwerpunkt des Kristalls, dar :

$$S_{ii}{}^{mm} = \overline{s_i{}^m{}^2}.$$

Das Matrixelement $S_{ik}{}^{mm} = \overline{s_i{}^m s_k{}^n}$ ist ein Mass für die statistische Abhängigkeit zwischen den Verschiebungskomponenten $s_i{}^m$ und $s_k{}^n$ in den Gitterpunkten $\boldsymbol{\nu}_m$ und $\boldsymbol{\nu}_n$. Es hängt nach (2.10) nur vom Abstand $\boldsymbol{\nu}_{m-n}$ dieser beiden Gitterpunkte ab, und zwar ist es eine gerade Funktion dieses Abstandes.

Dividiert man $S_{ik}{}^{mn}$ durch die Schwankungsquadrate, so erhält man den „Korrelationskoeffizienten"

$$\gamma_{ik}{}^{mn} = \frac{\overline{s_i{}^m s_k{}^n}}{\sqrt{(\overline{s_i{}^m{}^2}\,\overline{s_k{}^n{}^2})}}. \qquad (2.12)$$

In kubischen Gittern gilt speziell

$$\gamma_{ik}{}^{mn} = \frac{\overline{s_i{}^m s_k{}^n}}{\overline{s_i{}^m{}^2}} = \frac{S_{ik}{}^{mn}}{S_{ii}{}^{mm}}. \qquad (2.13)$$

Es ist stets $|\gamma_{ik}{}^{mn}| < 1$ für $\boldsymbol{m} \neq \boldsymbol{n}$, und weiter $\gamma_{ii}{}^{mm} = 1$. Um die Bedeutung des Korrelationskoeffizienten einzusehen, betrachte man den Abstand $\boldsymbol{R}_{m-n}$ zweier in den Gitterpunkten $\boldsymbol{\nu}_m$ und $\boldsymbol{\nu}_n$ gelegener Atome

$$\boldsymbol{R}_{m-n} = (\boldsymbol{s}_m - \boldsymbol{s}_n + \boldsymbol{\nu}_{m-n})$$

und bilde das Schwankungsquadrat der i-Komponente :

$$\overline{(X_i{}^{m-n} - x_i{}^{m-n})^2} = 2\overline{s_i{}^m{}^2}(1 - \gamma_{ii}{}^{mn}). \qquad (2.14)$$

Für $\gamma_{ii}{}^{mn} = 0$ ist also die mittlere quadratische Abweichung des Atomabstandes vom Gleichgewichtsabstand gleich dem doppelten Verschiebungsquadrat; die beiden Atome sind dann statistisch unabhängig voneinander. Für $\gamma_{ii}{}^{mn} > 0$ wird das Schwankungsquadrat kleiner als dieser Wert; das bedeutet, dass die i-Komponenten der beiden Atome im Mittel etwas in Phase schwingen. Für $\gamma_{ii}{}^{mn} < 0$ wird das Schwankungsquadrat noch grösser als bei statistischer Unabhängigkeit. Die Atome schwingen dann in Gegenphase. Für das Schwankungsquadrat des Atomabstandes selbst erhält man

$$\overline{(\boldsymbol{R}_{m-n} - \boldsymbol{\nu}_{m-n})^2} = 2\overline{\boldsymbol{s}_m{}^2}(1 - \tfrac{1}{3} \sum_{i=1}^{3} \gamma_{ii}{}^{mn}). \qquad (2.15)$$

Der Korrelationskoeffizient $\gamma_{ii}{}^{mn}$ bestimmt also gemäss den Gleichungen (2.14) und (2.15) das Schwankungsquadrat des Atomabstandes.

Es ist leicht zu zeigen, dass der Korrelationskoeffizient noch eine andere Bedeutung hat : Beobachtet man die Verteilung der Verschiebungen vom Schwerpunkt des Periodizitätsvolumens aus, und weiss man bereits, dass das Atom in $\boldsymbol{\nu}_m$ eine Verschiebungskomponente $s_i{}^m$ besitzt, so hat die Komponente $s_k{}^n$ des in $\boldsymbol{\nu}_n$ befindlichen Atoms im Mittel den Wert

$$\overline{s_k{}^n} = \gamma_{ik}{}^{mn} s_i{}^m.$$

Im kubisch flächenzentrierten Gitter besitzt jedes Atom 12 nächste und 6 übernächste Nachbarn (Abb. 2) Es gibt also insgesamt 6×12 Korrelationskoeffizienten zwischen den nächsten und 6×6 zwischen den übernächsten Nachbarn. Infolge der Gittersymmetrie sind jedoch die meisten dieser Werte untereinander gleich, während eine ganze Anzahl auch null wird. Durch einfache Symmetrie-Betrachtungen gewinnt man leicht das in Tabelle 1 angegebene Schema. Dort bedeutet

Tabelle 1. *Die Korrelationskoeffizienten* $\gamma_{ik}{}^{m\,m+n}$ *für die* 12 *nächsten und* 6 *übernächsten Nachbarn im kubisch flächenzentrierten Gitter*

Es ist $\gamma_{ik}{}^{m\,m+n} = \overline{s_i{}^m\, s_k{}^{m+n}}$ der Korrelationskoeffizient zwischen den in $\boldsymbol{\nu}^m$ und $\boldsymbol{\nu}^{m+n}$ befindlichen Atomen.

$\gamma_{ik}{}^{m\,m+n}$ für nächste Nachbarn

	x_1, x_2-Ebene	x_1, x_3-Ebene	x_2, x_3-Ebene
$\boldsymbol{n}$	$\{0, 0, 1\}$ $\{0, 0, \bar{1}\}$ $\{1, \bar{1}, 0\}$ $\{\bar{1}, 1, 0\}$	$\{0, 1, 0\}$ $\{0, \bar{1}, 0\}$ $\{1, 0, \bar{1}\}$ $\{\bar{1}, 0, 1\}$	$\{1, 0, 0\}$ $\{\bar{1}, 0, 0\}$ $\{0, 1, \bar{1}\}$ $\{0, \bar{1}, 1\}$
(γ)	$\begin{pmatrix} \alpha & \phi & 0 \\ \phi & \alpha & 0 \\ 0 & 0 & \beta \end{pmatrix}$	$\begin{pmatrix} \alpha & 0 & \phi \\ 0 & \beta & 0 \\ \phi & 0 & \alpha \end{pmatrix}$	$\begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & \phi \\ 0 & \phi & \alpha \end{pmatrix}$

$\gamma_{ik}{}^{m\,m+n}$ für übernächate Nachburn

	x_1-Achse	x_2-Achse	x_3-Achse
$\boldsymbol{n}$	$\{\bar{1}, 1, 1\}$ $\{1, \bar{1}, \bar{1}\}$	$\{1, \bar{1}, 1\}$ $\{\bar{1}, 1, \bar{1}\}$	$\{1, 1, \bar{1}\}$ $\{\bar{1}, \bar{1}, 1\}$
(γ)	$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix}$	$\begin{pmatrix} \mu & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix}$	$\begin{pmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \lambda \end{pmatrix}$

$\boldsymbol{am}$ die Komponenten des Verschiebungsvektors $\boldsymbol{\nu}^m = (A)\boldsymbol{m}$ zum nächsten bezw. übernächsten Nachbarn eines Atoms, bezogen auf die primitive Basis der a_i (Abb. 2). Es gibt also für nächste bezw. übernächste Nachbarn nur drei bezw. zwei unabhängige Werte α, β, ϕ bzw. λ, μ. Da das Matrixelement (2.11) nur vom Abstand $\boldsymbol{\nu}^{m-n}$ abhängt, liegt es nahe, eine stetige „Korrelationsfunktion" $S_{ik}(\boldsymbol{\nu})$ einzuführen, die an den Gitterpunkten $\boldsymbol{\nu}^{m-n}$ mit $S_{ik}{}^{mn}$ übereinstimmt und die Zwischenwerte stetig interpoliert. Damit wird also den Betrachtungen ein Kontinuum zugrundegelegt, in dem an jedem Raumpunkt Verschiebungen $\boldsymbol{s}(\boldsymbol{\nu})$ und Korrelationen zwischen diesen Verschiebungen definiert sind. Berechnet man dieses Verschiebungsfeld $\boldsymbol{s}(\boldsymbol{\nu})$ in Debyescher Näherung, so stimmt es mit dem eines elastischen Kontinuums überein. Verwendet man jedoch das richtige Gitterspektrum, so handelt es sich um ein fiktives Kontinuum, dessen Verschiebungsfeld nur an den Punkten des Raumgitters mit dem des Kristalls übereinstimmt.

Die Korrelationsfunktion erhält man, indem man in (2.10) $\boldsymbol{\nu}^{m-n}$ durch $\boldsymbol{\nu}$ ersetzt :

$$S_{ik}(\boldsymbol{\nu}) = \frac{2}{N} \sum_{k_1>0} s_{ik}{}^{\boldsymbol{k}} \cos \boldsymbol{k\nu}. \qquad (2.10a)$$

Das ist gerade die Fourier-Entwicklung der Korrelationsfunktion. Die Fourier-Koeffizienten hängen nach (2.11) vom Frequenzspektrum und von der Richtung der Polarisationsvektoren ab. Bei bekannter Korrelationsfunktion lässt sich also, wenn auch der Verlauf der Polarisationsvektoren $\boldsymbol{e}_{\boldsymbol{k}\lambda}$ bekannt ist, durch Umkehrung von (2.10) und (2.11) der Dispersionsverlauf $\omega_{\boldsymbol{k}\lambda} = \omega_\lambda(\boldsymbol{k})$ bestimmen.* Dividiert man die Korrelationsfunktion durch die Schwankungsquadrate, so erhält man den Korrelationskoeffizienten

$$\gamma_{ik}(\boldsymbol{\nu}) = \frac{S_{ik}(\boldsymbol{\nu})}{\sqrt{S_{ii}(0) S_{kk}(0)}}. \qquad (2.12a)$$

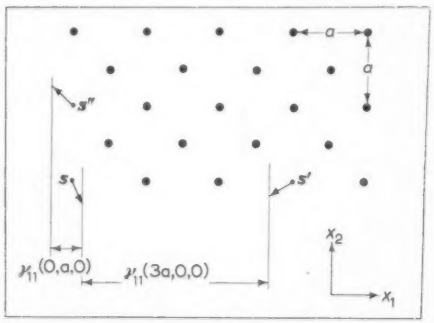


ABB. 3. Zur Bedeutung der Korrelationskoeffizienten $\gamma_{11}(\boldsymbol{\nu})$ der Verschiebungen. Es ist $\gamma_{11}(3a, 0, 0) = \overline{s_1 s_1'}$ und $\gamma_{11}(0, a, 0) = \overline{s_1 s_1''}$. Die Punkte bezeichnen die Ruhelagen der Atome im Gitter, die Endpunkte der Pfeile beliebige simultane Atomlagen.

* Ein Beispiel hierfür liefert eine Arbeit von Curien, der mittels Röntgen-Streuung die Grösse

$$\sum_{ik=1}^{3} k_i k_k s_{ik}{}^{\boldsymbol{k}}$$

in Abhängigkeit von der Differenz $\boldsymbol{k}$ der Ausbreitungsvektoren des einfallenden und des gestreuten Röntgen-Strahles für α-Fe gemessen und daraus den Dispersionsverlauf $\omega_\lambda(\boldsymbol{k})$ berechnet hat.[5]

Für kubische Gitter gilt überdies

$$\gamma_{ik}(\boldsymbol{\nu}) = \frac{S_{ik}(\boldsymbol{\nu})}{S_{ii}(0)}. \tag{2.13a}$$

Um die Indizierung der Korrelationskoeffizienten zu veranschaulichen, sind in Abb. 3 einige Verschiebungen in der x_1, x_2-Ebene des kubisch flächenzentrierten Gitters und die zwischen deren Komponenten wirksamen Korrelationkoeffizienten eingezeichnet. Es ist zu beachten, dass in $\gamma_{ik}(x_1, x_2, x_3)$ die Komponenten des Abstandsvektors auf die orthogonale Basis der $\boldsymbol{e}_i$ zu beziehen sind, während in $\gamma_{ik}{}^{mm+n}$ die m_i, n_i die entsprechenden Komponenten in dem primitiven Basissystem der $\boldsymbol{a}_i$ bedeuten, nämlich $x_i{}^m = (\boldsymbol{\nu}\boldsymbol{m}, \boldsymbol{e}_i)$, $(a^2/2)m_i = (\boldsymbol{\nu}\boldsymbol{m}, \boldsymbol{a}_i)$ (Abb. 2).

In Abb. 4 ist als Beispiel der Verlauf des Korrelationskoeffizienten $\gamma_{11}(\boldsymbol{\nu})$ im kubisch flächenzentrierten Gitter für die beiden Grenzfälle $T = 0$ ($\epsilon_{\boldsymbol{k}\lambda} = \frac{1}{2}\hbar\omega_{\boldsymbol{k}\lambda}$) und $kT \gg \hbar\omega_g$ ($\epsilon_{\boldsymbol{k}\lambda} = kT$, ω_g = Grenzfrequenz), also für die Nullpunktsschwingungen und für die klassische Wärmebewegung, dargestellt. Die Rechnung wurde mittels Debyescher Näherung für den isotropen Fall ausgeführt und ist im Anhang *a* erläntert.

Aufgetragen ist in Abb. 4 der Verlauf von $\gamma_{11}(\boldsymbol{\nu})$ längs der x_1- und x_2-Achse, die mit den Kanten des Elementarwürfels zusammenfallen (Abb. 2). Als Abzisse ist das Verhältnis x_i/a gewählt ($i = 1, 2$), wobei a die Gitterkonstante ist. $\gamma_{11}(a, 0, 0)$ und $\gamma_{11}(0, a, 0)$ sind also die beiden Korrelationskoeffizienten zwischen den übernächsten Nachbarn und senkrecht dazu (Abb. 3).

Wie ersichtlich, nehmen die Korrelationskoeffizienten bei hoher Temperatur viel langsamer mit der Entfernung ab als bei $T = 0$. Das ist verständlich, da bei hoher Temperatur die langen Wellen, die ja bei grossen Abständen vorwiegend für die Korrelation verantwortlich sind, im Verhältnis viel stärker angeregt sind als bei den Nullpunktsschwingungen bei $T = 0$. Ausserdem fällt die Korrelation senkrecht zu den Verschiebungen in beiden Fällen jeweils schneller ab als in Richtung der Verschiebungen. Das ist eine Folge der Verschiedenheit der zwischen zwei Atomen wirkenden Kräfte, wenn man das eine einmal in Richtung ihrer Verbindungslinie und sodann senkrecht dazu auslenkt. Die Reichweite der Korrelation is von der Grössenordunng einer Gitterkonstanten.

Der Kurvenverlauf als Ganzes gibt nach (2.12a) auch unmittelbar das relative Schwankungsquadrat der Komponenten des Atomabstandes an, nämlich

$$1-\gamma_{ii}(\boldsymbol{\nu}) = \frac{\overline{(X_i(\boldsymbol{\nu})-x_i)^2}}{2\overline{s_i{}^2}}.$$

Dieses Schwankungsquadrat verschwindet natürlich für $\boldsymbol{\nu} = 0$ und besitzt für nächste Nachbarn seinen kleinsten Wert. Für grosse Atomabstände ($|\boldsymbol{\nu}| \gg a$) beträgt dann der Wert praktisch immer eins.

Da bei der mittels Debyescher Näherung durchgeführten Berechnung der Kurven auch Isotropie vorausgesetzt wurde, hängt der gezeichnete Verlauf nicht von der Orientierung des Koordinatensystems im Kristall ab. Legt man die

Tabelle 2. $\gamma_{11}(\boldsymbol{\nu})$ für nächste und übernächste Nachbarn in Debyescher Näherung

$\boldsymbol{\nu}$			$\gamma_{11}(\boldsymbol{\nu})$	
			$T = 0$	$kT \gg \hbar\omega_g$
$(x_1, 0, 0)$	(Korrelation in Richtung der Verschiebungen)	n.N.$(x_1 = a/\sqrt{2})$	0,2	0,5
		ü.n.N.$(x_1 = a)$	0,08	0,35
$(0, x_2, 0)$	(Korrelation senkrecht zur Richtung der Verschiebungen)	n.N.$(x_2 = a/\sqrt{2})$	0,09	0,34
		ü.n.N.$(x_2 = a)$	−0,05	0,2

x_1-Achse nicht in die Kante, sondern in die Flächendiagonale des Elementarwürfels, so ergibt sich also für $\gamma_{11}(\boldsymbol{\nu})$ der gleiche Kurvenverlauf. Für $x_1 = a/\sqrt{2}$ erhält man so den Korrelationskoeffizienten zwischen nächsten Nachbarn, der in Abb. 4 gestrichelt eingetragen ist.

Auf diese Weise lassen sich Abb. 4 die Werte der Korrelationskoeffizienten für die nächsten und übernächsten Nachbarn entnehmen, die in Tabelle 2 zusammengestellt sind. Diese Zahlenwerte vermitteln eine ungefähre Vorstellung von den Korrelationsverhältnissen in der Umgebung eines Atoms. Augenscheinlich spielt z.B. auch die Korrelation zwischen übernächsten Nachbarn noch eine merkliche Rolle. Aus den Werten von Tabelle 2 erhält man auch sofort, wie im Anhang (a) angegeben wird, die Zahlenwerte für die in Tabelle 1 auftretenden Korrelationskoeffizienten. Diese Werte sind in Tabelle 3 eingetragen. Die Debyesche Näherung liefert also keine Korrelation zwischen verschiedenen Verschiebungskomponenten, es ist $\gamma_{ik} = \gamma_{ii}\,\delta_{ik}$.

Tabelle 3. *Die Zahlenwerte der Korrelationskoeffizienten aus Tabelle* 1 *in Debyescher Näherung für die Nullpunktsschwingungen und für die klassische Wärmebewegung*

		$T=0$	$kT \gg \hbar\omega_g$
n.N.	α	0,15	0,42
	β	0,09	0,34
	ϕ	0	0
ü.n.N.	λ	0,08	0,35
	μ	−0,05	0,2

Offensichtlich geben also die mittels Debyescher Näherung für den isotropen Fall berechneten Korrelationsfunktionen die Verhältnisse im Kristall nur angenähert wieder. Das beruht darauf, dass hierbei der Einfluss der kurzen Wellen nur ungenügend berücksichtigt wird.

(c) *Verzerrungen und Spannungen*

Für die dynamischen Vorgänge im Kristall sind nicht die Verschiebungen, sondern die Verzerrungen

$$\epsilon_{ik}(\boldsymbol{\nu}) = \frac{1}{2}\left(\frac{\partial s_i(\boldsymbol{\nu})}{\partial x_k} + \frac{\partial s_k(\boldsymbol{\nu})}{\partial x_i}\right) \tag{2.16}$$

massgebend, die aber zunächst nur im Kontinuum definiert sind. Um den Anschluss an die elastische Theorie herzustellen, werden, wie bereits oben bei der Korrelationsfunktion durchgeführt, die Verschiebungen $s_i{}^m$ in ein Verschiebungsfeld $s_i(\boldsymbol{\nu})$ eingebettet, das an den Gitterpunkten $\boldsymbol{\nu}m$ mit den Verschiebungen des dortigen Atoms zusammenfällt : $s_i(\boldsymbol{\nu}m) = s_i{}^m$ Nach (2.1) und (2.2b) ist dieses Feld gegeben durch

$$s_i(\boldsymbol{\nu}) = \sqrt{\left(\frac{2}{N}\right)} \sum_{\substack{k_1>0 \\ \lambda}} e_{\lambda i}{}^k \{q_{k\lambda} \cos \boldsymbol{k\nu} + q_{-\boldsymbol{k}\lambda} \sin \boldsymbol{k\nu}\}. \tag{2.17}$$

Die Werte des daraus nach (2.16) abgeleiteten Verzerrungsfeldes in den Gitterpunkten $\boldsymbol{\nu}m$ sollen mit $\epsilon_{ik}{}^m$ bezeichnet werden : $\epsilon_{ik}(\boldsymbol{\nu}m) = \epsilon_{ik}{}^m$ Damit sind nun auch im Gitter Verzerrungen eingeführt, die sich wegen (2.2b) schreiben lassen :

$$\epsilon_{ik}{}^m = \sum_{\boldsymbol{k}} \alpha_{\boldsymbol{m}}{}^{\boldsymbol{k}} v_{ik}{}^{\boldsymbol{k}}. \tag{2.18}$$

Dabei sind

$$v_{ik}{}^{\boldsymbol{k}} = \tfrac{1}{2} \sum_{\lambda} (e_{\lambda i}{}^k k_k + e_{\lambda k}{}^k k_i) q_{-\boldsymbol{k}\lambda} \tag{2.19}$$

neue Normalkoordinaten.

Von den $6N$ Variablen $\varepsilon_{ik}{}^m$ sind nur $3N-3$ unabhängig, denn in jedem der N Gitterpunkte des Periodizitätsvolumens müssen 3 Kompatibilitätsbedingungen und ausserdem 3 Bedingungen, welche die Konstanz des Schwerpunktes ausdrücken, erfüllt sein. Will man das statistische Verhalten des Kristalls in diesen $6N$ Variablen beschreiben, so ergibt sich analog zu (2.9) eine Gauss-Verteilung, in der aber noch $3N+3$ δ-Funktionen zur Berücksichtigung der Nebenbedingungen stehen. Stattdessen kann man aber auch eine Verteilungsfunktion mit einer Auswahl von nur $3N-3$ unabhängigen Variablen verwenden. In beiden Fällen lässt sich aber die im Exponenten stehende Matrix nicht explizit angeben, weil die $\varepsilon_{ik}{}^m$ von den $q_{k\lambda}$ nach (2.18) mit (2.19) nicht mehr durch eine orthogonale Transformation abhängen. Dagegen hängt bei festem i, k $\varepsilon_{ik}{}^m$ nach (2.18) von $v_{ik}{}^{\boldsymbol{k}}$ durch eine orthogonale Transformation ab.

Da jedes $v_{ik}{}^{\boldsymbol{k}}$ von den $q_{\boldsymbol{k}\lambda}$ linear abhängt,

besitzt es eine Gauss-Verteilung,* für deren Schwankungsquadrat man aus (2.18) mit (2.11) sofort erhält :

$$\overline{v_{ik}{}^{k^2}} = \overline{v_{ik-k}{}^2} = \epsilon_{ik}{}^k$$

$$= \tfrac{1}{4}(k_i{}^2 s_{kk}{}^k + k_k{}^2 s_{ii}{}^k + 2k_i k_k s_{ik}{}^k). \qquad (2.20a)$$

Bei festgehaltenem Schwerpunkt des Kristalls ist (2.20a) durch den Wert für $\boldsymbol{k} = 0$ zu ergänzen :

$$\overline{v_{ik}{}^{02}} = \epsilon_{ik}{}^0 = 0. \qquad (2.20b)$$

Für festes $i,\boldsymbol{k}$ sind die N-Grössen $v_{ik}{}^k$ statistisch unabhängig, sodass ihre Verteilung das Produkt der einzelnen Gauss-Verteilungen ist. Setzt man darin (2.18) ein, so erhält man genau wie bei den Verschiebungen die Verteilung für die N Verzerrungen $\epsilon_{ik}{}^m$ bei festem $i,\boldsymbol{k}$:

$$w(\ldots \epsilon_{ik}{}^m \ldots) \ldots d\epsilon_{ik}{}^m \ldots = \frac{1}{(2\pi)^{N-1/2}|E'_{ik}|^{1/2}} \times$$

$$\times \exp(-\tfrac{1}{2} \sum_{mn} \hat{E}_{ik}{}^{mn} \epsilon_{ik}{}^m \epsilon_{ik}{}^n) \delta(\sum_n \epsilon_{ik}{}^n) \ldots d\epsilon_{ik}{}^m \ldots \qquad (2.21)$$

Die Korrelationsmatrix wird :

$$E'_{ik}{}^{mn} = \overline{\epsilon_{ik}{}^m \epsilon_{ik}{}^n} = \frac{2}{N} \sum_{k_1>0} \epsilon_{ik}{}^k \cos \boldsymbol{k}\nu_{m-n}. \qquad (2.22a)$$

Weiter ist

$$\hat{E}_{ik}{}^{mn} = \frac{2}{N} \sum_{k_1>0} \frac{1}{\epsilon_{ik}{}^k} \cos \boldsymbol{k}\nu_{m-n} \qquad (2.22b)$$

und

$$|E_{ik}'| = \prod_{\boldsymbol{k}}{}' \epsilon_{ik}{}^k. \qquad (2.22c)$$

Wieder wird eine stetige Korrelationsfunktion $E_{ik}(\boldsymbol{\nu})$ eingeführt, die $E_{ik}{}^{mn}$ interpoliert, und dann auf 1 normiert :

$$e_{ik}(\boldsymbol{\nu}) = \frac{E_{ik}(\boldsymbol{\nu})}{E_{ii}(0)}.$$

* Hängt eine Grösse von mehreren Variablen, die unabhängig gaussisch verteilt sind, linear ab, so besitzt diese Grösse ebenfalls eine Gauss-Verteilung. Siehe z.B. UHLENBECK.[6]

Abb. 5 zeigt als Beispiel den Verlauf von $e_{12}(x_1, 0, 0)$ und $e_{12}(0, 0, x_3)$ für $T = 0$ und $kT \gg \hbar\omega_g$, der im Anhang (a) mittels Debyescher Näherung für das kubisch flächenzentrierte

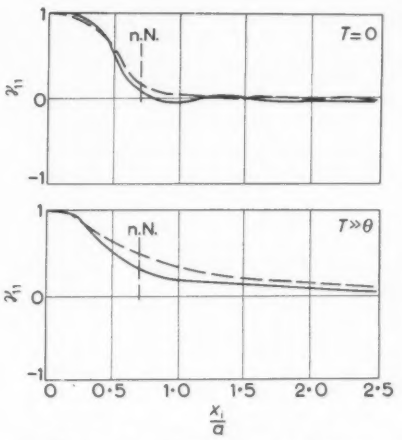


ABB. 4. Verlauf des Korrelationskoeffizienten der Verschiebungen $s_1{}^n$. (- - - - : $\gamma_{11}(x_1, 0, 0)$; — : $\gamma_{11}(0, x_2, 0)$).

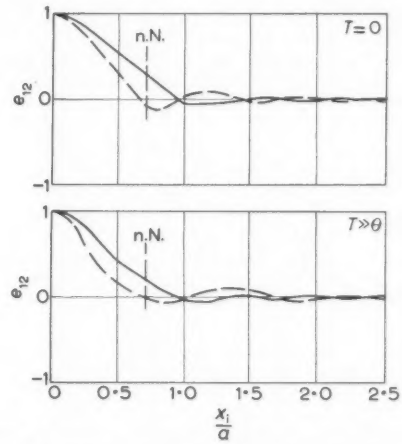


ABB. 5. Verlauf der Korrelationskoeffizienten der Verzerrungen $\varepsilon_{12}{}^n$ bezw. der Schubspannungen $\tau_{12}{}^n$. (- - - - - - : $e_{12}(x_1, 0, 0)$; — : $e_{12}(0, 0, x_3)$).

Gitter berechnet wird. Die beiden Kurvenpaare stellen also den Korrelationsverlauf der Verzerrung $\epsilon_{12}{}^m$ innerhalb der x_1, x_2-Ebene und senkrecht dazu dar. Man erkennt, dass der Korrelationsverlauf senkrecht zur x_1, x_2-Ebene praktisch temperaturunabhängig ist; ausserdem ist die

Korrelation hier im wesentlichen immer positiv. Der Korrelationsverlauf innerhalb der x_1, x_2-Ebene ist dagegen etwas temperaturabhängig, derart, dass bei $T = 0$ der Bereich negativer Korrelation stärker ins Gewicht fällt.

Dass die Korrelation der Verzerrungen ε_{12} zwischen benachbarten Atomen innerhalb der x_1, x_2-Ebene negativ, senkrecht dazu aber positiv ist, ersieht man aus Abb. 6. Dort sind -der Einfachheit halber nur für das kubisch primitive Gitter- einige mögliche Schwingungszustände dargestellt. Die Verzerrungen ε_{12} bestimmen jeweils den Winkel α zwischen einem Atom und seinen beiden Nachbarn in $+x_1$- und $+x_2$-Richtung; es ist nämlich $\alpha = (\pi/2) - 2\varepsilon_{12}$. Durch den Wert von ε_{12} des Atoms (1) ist also nur dieser Winkel α, aber nicht die Lage der Atome (2) und (3) bestimmt. Für einen festen (positiven) Wert ε_{12} in (1) ist die wahrscheinlichste Lage der Atome (2) und (3) diejenige, bei der die Winkelhalbierends von α auch die Winkelhalbierende des positiven Quadranten ist; diese Lage ist in Abb. 6 eingezeichnet. In der linken Abb. 6(a) haben die Winkel α' und α'' in (2) und (3) den gleichen Wert wie α in (1); dies ist jedoch ein äusserst unwahrscheinlicher Schwingungszustand, weil die dazu erfordelichen Verschiebungskomponenten s_2' von (2), s_1'' von (3) und $s_1''' = s_2'''$ von (4) sehr gross sind. Der wahrscheinlichste Schwingungszustand wird etwa so wie die rechte Abb. 6(a) aussehen : die Verschiebungskomponenten s_2', s_1'', $s_1''' = s_2'''$ sind so klein, dass die Winkel α'' und α' grösser als $\pi/2$, also ε' und ε'' negativ sind. Das bedeutet aber, dass die Korrelationskoeffizienten γ'' zwischen ε und ε'' und γ' zwischen ε und ε' negativ sind. Im Gegensatz hierzu ist der wahrscheinlichste Wert von α' des auf der x_3-Achse gelegenen Atoms (2) für einen fest vorgegebenen positiven Wert α in (1) kleiner als $\pi/2$ (Abb. 6b); mit ε ist also auch ε' positiv. Das bedeutet, dass der Korrelationskoeffizient γ' zwischen ε und ε' positiv ist.

Die unterschiedliche Korrelation von ε_{12} innerhalb und senkrecht zur x_1, x_2-Ebene wird später bei der Erklärung der Volumenabhängigkeit des Volumenmittels von ε_{12} bezw. $\tau_{12} = 2c_{44}\varepsilon_{12}$ von Bedeutung sein.

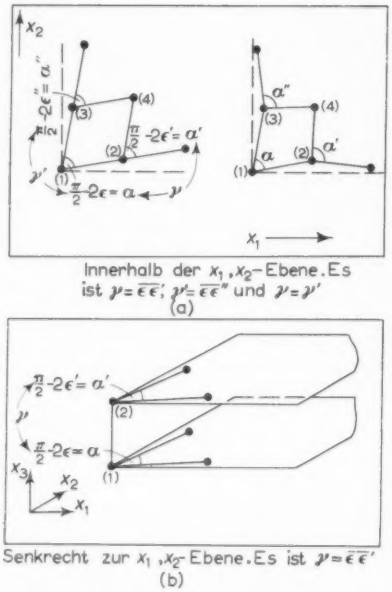

Innerhalb der x_1, x_2-Ebene. Es ist $\gamma = \overline{\varepsilon\varepsilon}$; $\gamma' = \overline{\varepsilon\varepsilon''}$ und $\gamma = \gamma'$
(a)
Senkrecht zur x_1, x_2-Ebene. Es ist $\gamma = \overline{\varepsilon\varepsilon'}$
(b)

Abb. 6. Zur Korrelation der Verzerrungen $\varepsilon = \varepsilon_{12}$ innerhalb der x_1, x_2- Ebene (a) und senkrecht dazu (b). Die Punkte bezeichnen simultane Atomlagen. Die -nicht eingezeichneten- Ruhelagen der Atome bilden ein kubisch primitives Gitter.

Nachdem in den Gitterpunkten Verzerrungen $\epsilon_{ik}{}^m$ definiert sind, lassen sich ganz entsprechend auch elastische Spannungen $\tau_{mn}{}^m$ einführen. Diese sind zunächst im elastischen Kontinuum mit dem Verschiebungsfeld $\boldsymbol{s}(\boldsymbol{\nu})$ gegeben durch :

$$\tau_{mn}(\boldsymbol{\nu}) = \sum_{ik=1}^{3} C_{mn,ik} \frac{\partial s_i(\boldsymbol{\nu})}{\partial x_k} \qquad (2.23)$$

$C_{mn,ik}$ ist der vierstufige Tensor der elastischen Konstanten. Verwendet man jetzt als Verschiebungsfeld $\boldsymbol{s}(\boldsymbol{\nu})$ wie oben das nichtelastische fiktive Kontinuum, dessen Verschiebungsfeld in den Punkten des Raumgitters mit dem des Kristalls übereinstimmt, so lassen sich analog wie bei den Verzerrungen $6N$ Spannungsvariable definieren durch

$$\tau_{mn}{}^m = \sum_k \alpha_m{}^k w_{mn}{}^k. \qquad (2.24)$$

Dabei treten bei festem m, n de N-Grössen $w_{mn}{}^k$ als neue Normalkoordinaten auf.

Sie sind unabhängig gaussisch verteilt und besitzen das Schwankungsquadrat

$$\overline{w_{mn}{}^{k2}} = \overline{w_{mn}{}^{-k2}} \equiv \tau_{mn}{}^k = \sum_{ijkl} C_{mn,ij} C_{mn,kl} k_j k_l s_{ik}{}^k \qquad (2.25)$$

Für die Verteilungsfunktion der $\tau_{mn}{}^m$ gelten die gleichen Bemerkungen wie für die symmetrischen Verzerrungen $\epsilon_{ik}{}^m$. Explizit angeben lässt sich allein eine Verteilung für N statistische Variable $\tau_{mn}{}^m$ bei festem m, n. Man erhält diese Verteilung sofort durch Einsetzen von (2.24) in die Verteilung der $w_{mn}{}^k$, wobei wieder die Orthogonalitä

von (2.24) zu beachten ist. Bei festgehaltenem Schwerpunkt ergibt sich somit :

$$w(\ldots\tau_{mn}{}^{m}\ldots)\ldots d\tau_{mn}{}^{m}\ldots = \frac{1}{(2\pi)^{N-1/2}|T'_{mn}|^{1/2}}\times$$

$$\times\exp(-\tfrac{1}{2}\sum_{mn}\hat{T}_{mn}{}^{mn}\tau_{mn}{}^{m}\tau_{mn}{}^{n})\delta(\sum_{n}\tau_{mn}{}^{n})\ldots d\tau_{mn}{}^{m}\ldots \tag{2.26}$$

$$T'_{mn}{}^{mn} = \overline{\tau_{mn}{}^{m}\tau_{mn}{}^{n}} = \frac{2}{N}\sum_{k_1>0}\tau_{mn}{}^{k}\cos \boldsymbol{k}\boldsymbol{\nu}_{m-n} \tag{2.27a}$$

$$\hat{T}_{mn}{}^{mn} = \frac{2}{N}\sum_{k_1>0}\frac{1}{\tau_{mn}{}^{k}}\cos \boldsymbol{k}\boldsymbol{\nu}_{m-n} \tag{2.27b}$$

$$|T'_{mn}| = \prod_{k}{}'\,\tau_{mn}{}^{k}. \tag{2.27c}$$

Im isotropen Fall ist*

$$C_{mn,ik} = c_{12}\delta_{mn}\delta_{ik}+c_{44}(\delta_{mi}\delta_{nk}+\delta_{ni}\delta_{mk}). \tag{2.28}$$

Dann folgt z.B. für $\tau_{12}{}^{m}$ aus (2.23) :

$$\tau_{12}{}^{m} = 2g\epsilon_{12}{}^{m}. \tag{2.29}$$

In diesem Fall sind die durch (2.26) und (2.27) für die $\tau_{12}{}^{m}$ gegebene Verteilung und deren Korrelationskoeffizient—bis auf den unwesentlichen Faktor $2c_{44} = 2g$—mit der Verteilung (2.21) und (2.22) für die symmetrischen Verzerrungen $\epsilon_{12}{}^{m}$ identisch. Auch der in Abb. 5 gezeichnete Korrelationsverlauf der symmetrischen Verzerrungen stimmt dann mit dem entsprechenden Verlauf der Schubspannungen überein.

Interessiert nicht die Verteilung der Spannungen in dem ganzen Periodizitätsvolumen, sondern nur in einem Teilvolumen, so erhält man diese Verteilung durch Integration über die übrigen Variablen. Es ergibt sich auf diese Weise wieder eine Gauss-Verteilung, wobei die im Exponenten auftretende Matrix die Reziproke derjenigen Untermatrix der Korrelationsmatrix ist, die aus den in der Verteilung explizit auftretenden Variablen gebildet wird. Besonders einfach ist der Fall von nur einer Variablen. Z.B. erhält man für die lokale Schubspannung $\tau_{12}{}^{m} = \tau_0$ in einem beliebigen Gitterpunkt eine Gauss-Verteilung, deren Schwankungsquadrat das Diagonalelement von (2.27a) ist; es besitzt in Debyescher Näherung für $T = 0$ und im klassischen Grenzfall hoher Temperatur nach Anhang (a) den Wert

$$\tau_0{}^2 = \begin{cases} 0{,}2g\dfrac{k\theta}{v_0} & (T = 0) \\ 0{,}47g\dfrac{kT}{v_0} & (T \gg \theta) \end{cases} \tag{2.30}$$

θ ist die Debye-Temperatur ($k\theta = \hbar\omega_g$, ω_g = transversale Grenzfrequenz), v_0 wieder das Volumen pro Atom.

3. STATISTISCHE BEGRÜNDUNG DER BECKERSCHEN FORMEL

Im folgenden soll nun die strenge Ableitung der Beckerschen Formel durch Berechnung der Verteilungsfunktion $w(\tau)$ für die mittlere Schubspannung $\tau = \overline{\tau_{12}}^{v}$ eines Probevolumens nach Gl. (1.3) durchgeführt werden. Dabei sollen insbesondere der Einfluss der Korrelation und der Art des Probevolumens untersucht werden.

(a) *Ohne Korrelation*

Den einfachsten Ausgangspunkt einer statistischen Begründung stellt die Annahme dar, dass die Schubspannungen in den n Gitterstellen des Probevolumens statistisch unabhängig verteilt sind. Die Verteilung für alle statistischen Variablen ist dann einfach das Produkt über n einzelne Gaussverteilungen,† für deren Schwankungsquadrat der in (2.30) angegebene, mittels Debyescher Näherung berechnete Wert $\overline{\tau_0{}^2}$ verwendet werden soll. Gemäss (1.2) hängt die mittlere Schubspannung τ von den $\tau_1, \ldots, \tau_n$ linear ab und besitzt daher ebenfalls eine Gauss-Verteilung‡

$$w(\tau)\,d\tau = \frac{1}{\sqrt{(2\pi\overline{\tau^2})}}\exp\left(-\frac{\tau^2}{2\overline{\tau^2}}\right)d\tau. \tag{3.1}$$

* Die c_{ik} sind die Voigtschen elastischen Konstanten. Die Isotropie-Bedingung lautet $2c_{44} = c_{11}-c_{12}$; $g = c_{44}$ ist der Schubmodul, $E = c_{44}(3c_{11}-4c_{44})/(c_{11}-c_{44})$ der Elastizitätsmodul.

† Man gelangt natürlich auch zur Verteilung (3.1), indem man im Exponenten der strengen Verteilung (2.26) alle Korrelationskoeffizienten gleich null setzt und nur die Diagonalglieder der quadratischen Form übrig behält.

‡ Hängt eine Grösse von mehreren Variablen, die unabhängig gaussisch verteilt sind, linear ab, so besitzt diese Grösse ebenfalls eine Gauss-Verteilung. Siehe z.B. UHLENBECK.[6]

Das Schwankungsquadrat beträgt wegen (1.2) :

$$\overline{\tau^2} = \frac{\overline{\tau_0^2}}{n} = \begin{cases} 0{,}2g\dfrac{k\theta}{v} & (T = 0) \\ 0{,}47g\dfrac{kT}{v}. & (T \gg \theta) \end{cases} \tag{3.2}$$

Die Gleichung (3.2) tritt jetzt an die Stelle der Beckerschen Formel (1.1b) und unterscheidet sich von dieser einmal durch einen Zahlenfaktor 2 im Schwankungsquadrat; ausserdem liefert die statistische Begründung auch für $T = 0$ einen von den Nullpunktsschwingungen herrührenden endlichen Wert für das Schwankungsquadrat.

Da die thermodynamisch begründete Formel (1.1b) ganz sicher für hohe Temperaturen und grosse Probevolumina stimmt, ist das Auftreten des falschen Zahlenfaktors 0,5 in (3.2b) offensichtlich auf die Vernachlässigung der Korrelation zurückzuführen. Es ist demnach zu erwarten, dass die Berücksichtigung der Korrelation für grosse Volumina v den richtigen Faktor 1 liefert.

(b) *Mit Korrelation*

Die streng gültige Formel erhält man durch Einsetzen der richtigen Verteilung, die aus (2.26) durch Integration über alle ausserhalb des Probevolumens gelegenen Atome entsteht, in Gleichung (1.3). Es ist jedoch garnicht erforderlich, (1.3) explizit auszurechnen, da von vornherein feststeht, dass die mittlere Schubspannung τ als von den $\tau_1, \ldots, \tau_n$ und damit auch von den Normalkoordinaten w^k linear abhängige Grösse eine Gauss-Verteilung besitzt. Es ist also nur das Schwankungsquadrat zu berechnen. Wegen (1.2) erhält man hierfür

$$\overline{\tau^2} = \frac{1}{n^2} \sum_{mn} \overline{\tau_m \tau_n}$$

wobei die Summation über die n Atome des Probevolumens zu erstrecken ist. Durch Einsetzen von (2.27a) und Vertauschung der Reihenfolge der Summationen ergibt sich :*

$$\overline{\tau^2} = \frac{2}{N} \sum_{k_i > 0} \tau_k D_k^2. \tag{3.3}$$

Die Summe ist über die halbe Brillouin-Zone zu erstrecken. Hierbei wurde zur Abkürzung

$$D_k^2 = \frac{1}{n^2} \sum_{m,n} \cos \boldsymbol{k}\boldsymbol{\nu}_{m-n} = \left| \frac{1}{n} \sum_{n} \exp(i\boldsymbol{k}\boldsymbol{\nu}_n) \right|^2 \tag{3.4}$$

gesetzt. Die Doppelsumme ist über die n Atome des Probevolumens zu erstrecken. Da das Auftreten von D_k^2 in (3.3) für die Volumenabhängigkeit von $\overline{\tau^2}$ verantwortlich ist, soll D_k^2 als „Volumenfaktor" bezeichnet werden.†

Wir berechnen D_k^2 für die drei Fälle eines Würfels der Kantenlänge L, eines Quaders, der aus zwei benachbarten quadratischen Netzebenenabschnitten der Kantenlänge L besteht, und eines quadratischen Prismas der Länge L, dessen Basis nur ein Atom enthält. Die Kanten sollen jeweils in die Achsenrichtungen des in Abschnitt 1 eingeführten Koordinatensystems x_1, x_2, x_3 fallen. Der flache Quader soll in der x_1, x_3-Ebene (Gleitebene), die Längsachse des Prismas in x_3-Richtung (senkrecht zur Gleitrichtung) liegen. In allen Fällen sei $L \gg a$. Dann erhält man, wie im Anhang (*b*) gezeigt wird, für die Volumenfaktoren die in Spalte 2 von Tabelle 4 angegebenen Ausdrücke.

Darin sind k_j die Komponenten des Wellenzahlvektors in dem angegebenen Koordinatensystem, l_j die Zahl der auf einem Abschnitt der Länge L auf der x_j -Achse gelegenen Atome.‡ Alle angegebenen Volumenfaktoren sind für $L \gg a$ gültige Näherungsausdrücke.§

* In τ und τ^k ist stets der Index (12), der sich auf das im Abschnitt 1 erklärte Koordinatensystem x_1, x_2, x_3 bezieht, fortgelassen. Beim kubisch- flächenzentrierten Gitter, das hier betrachtet wird, fällt die x_1-, x_2-, x_3-, Achse in die (1$\bar{1}$0-, (111)-, (1$\bar{1}$2)- Richtung.

† D_k^2 ist auch der bei der Röntgen-Streuung am Probevolumen auftretende Strukturfaktor, also Gitter-Strukturfaktor multipliziert mit dem Formfaktor des Probevolumens.

‡ Es ist also $L = l_1(a/\sqrt{2}) = l_2(a/\sqrt{3}) = l_3(\sqrt{3}a/\sqrt{2})$; siehe hierzu Abb. 8.

§ Ersetzt man die Summation durch Integration, so ergeben sich für beliebiges L entsprechende Ausdrücke, wo nur im Nenner statt des sin jeweils das Argument steht. Beide Gruppen von Ausdrücken lassen sich, wie im folgenden für die in Tabelle 4 angegebenen Formen gezeigt wird, durch eine Exponentialfunktion approximieren. Das bedeutet, dass für $L \gg a$ die Gittersumme durch ein Integral ersetzt werden darf.

Die in den Volumenfaktoren auftretende Funktion

$$u = \left(\frac{\sin lx}{l \sin x}\right)^2 \tag{3.5}$$

wird bei der später auszuführenden Summation über die $\boldsymbol{k}$-Werte der 1. Brillouin-Zone in dem Intervall von 0 bis $\pi/2$ benötigt. Ihre Rolle besteht also darin, dass sie die infolge der Mittelung über das Probevolumen fortfallenden kurzen Wellen abschneidet. Die Kurvenform interessiert also im einzelnen garnicht, und es ist daher statthaft, im Hinblick auf die spätere Auswertung der Summe (3.3) durch Integration die Funktion (3.5) durch eine flächengleiche Gauss-Funktion zu ersetzen, deren Funktionswert in $x = 0$ ausserdem mit dem von (3.5) übereinstimmt.*

$$u = \left(\frac{\sin lx}{l \sin x}\right)^2 \to \exp\left(-\frac{l^2}{\pi} x^2\right).$$

Ausserdem soll dann später die Integration über die betreffenden k_j statt bis zur Grenze der 1. Brillouin-Zone bis unendlich erstreckt werden; das bedeutet, dass hier die Integration über x bis unendlich zu führen ist.

Schliesslich soll noch der $\cos^2$-Term im Volumenfaktor des Prismas einfach durch den Mittelwert 1/2 ersetzt werden.

Es muss betont werden, dass diese vereinfachenden Annahmen das Ergebnis natürlich beeinflussen. Durch die Exponentialfunktion werden die kurzen Wellen schärfer abgeschnitten, als durch die strengen Funktionen (3.5). Im Integranden tritt aber später ausser der Abschneidefunktion im wesentlichen noch ein Faktor k bzw. k^2 für tiefe bzw. hohe Temperaturen auf, der gerade den Beitrag der kurzen Wellen mit einem grossen Gewicht versieht. Man wird also erwarten, dass das im folgenden berechnete Ergebnis etwas kleiner als der richtige Wert ist. Jede strenge Berechnung unter Verwendung der Funktionen (3.5) ist rechnerisch zu mühsam.

* Es gilt nämlich

$$\int_0^{\pi/2} \left(\frac{\sin lx}{l \sin x}\right)^2 dx = \int_0^{\infty} \exp\left(-\frac{l^2}{\pi^2} x^2\right) dx = \frac{\pi}{2l}.$$

Als Volumenfaktoren sollen also endgültig die Ausdrücke verwendet werden, die in Spalte 3 von Tabelle 4 angegeben sind. Darin bedeutet $k = |\boldsymbol{k}|$.

Es soll im folgenden das Schwankungsquadrat für den Fall eines isotropen Mediums in Debyescher Näherung berechnet werden. Dazu ist in (3.3) der Wert für τ^k einzusetzen, für den man in den beiden Grenzfällen $T = 0$ ($\epsilon_{k\lambda} = \frac{1}{2}\hbar c_\lambda k$) und $kT \gg \hbar\omega_g$ ($\epsilon_{k\lambda} = kT$, ω_g = transversale Grenzfrequenz) aus (2.25) in Verbindung mit (2.2a), (2.5) und (2.28) erhält :

$$\tau^k = \begin{cases} \dfrac{\hbar}{g^2 2mc_2}\left\{4(\delta-1)\dfrac{k_1^2k_2^2}{k^3} + \dfrac{k_1^2+k_2^2}{k}\right\} & (T=0) \\ \dfrac{kT}{g^2 mc_2^2}\left\{4(\delta^2-1)\dfrac{k_1^2k_2^2}{k^4} + \dfrac{k_1^2+k_2^2}{k^2}\right\} & (kT \gg \hbar\omega_g) \end{cases} \tag{3.6}$$

δ ist das Verhältnis der Schallgeschwindigkeiten ($\delta = c_2/c_1$) und hat für isotrope Metalle, wie W und Al, den Wert 1/2; dieser Wert soll fortan benutzt werden.

Zur Berechnung von (3.3) wird die Summation durch eine Integration ersetzt; diese ist über die 1. Brillouin-Zone zu erstrecken, deren Volumen beim kubisch flächenzentrierten Gitter $v_{\mathrm{Br}} = 4(2\pi/a)^3$ beträgt. Man hat somit

$$\overline{\tau^2} = \frac{1}{v_{\mathrm{Br}}} \iiint \tau_k D_k^2 \, d\boldsymbol{k} \tag{3.7}$$

für die beiden Werte (3.6) für τ^k und die drei in Spalte 3 von Tabelle 4 angegebenen Volumenfaktoren D_k^2 zu berechnen. Diese Rechnung ist im Anhang c erläutert und liefert die in Spalte 6 der Tabelle 4 eingetragenen Werte. Die angegebenen Ausdrücke sind jeweils nur das erste Glied einer asymptotischen Entwicklung nach Potenzen von (a/L); da nach Voraussetzung $L \gg a$ ist, sind die höheren Potenzen fortgelassen.

Es ist zu beachten, dass die Gültigkeitsgrenze für die auftretenden klassischen Werte beim Quader und Prisma durch die Debye-Temperatur θ ($k\theta = \hbar\omega_g$) gegeben ist, also durch die Bedingung $T \gg \theta$; beim Würfel werden jedoch durch den Volumenfaktor sämtliche Wellen bereits bei einer Frequenz $\omega_g^* = \omega_g(a/L)$ abgeschnitten, sodass

Tabelle 4. Zur Berechnung des Schwankungsquadrates $\overline{\tau^2}$ nach (3.7)

Spalte 2 enthält die strengen, Spalte 3 die angenäherteu Volumenfaktoren (3.4). Spalte 6 and 7 enthält die mit den Werten aus Spalte 3 berechneten (mit v/g multiplizierten) Schwankungsquadrate in den jeweils angegebenen Näherungen. Die thermodynamische Formel (1.1b) liefert für $v/g\overline{\tau^2}$ den Wert kT.

1. Volumen	2. $D_{\boldsymbol{k}}{}^2$ (streng)	3. $D_{\boldsymbol{k}}{}^2$ (angenähert)	4. Grenzfall	5. $v/g \cdot \overline{\tau^2}$ Ohne Korrelation in Deb. Näherung	6. $v/g \cdot \overline{\tau^2}$ Mit Korrelation in Deb. Näherung	7. $v/g \cdot \overline{\tau^2}$ Mit Korrelation und Gitterspektrum
Würfel : $V = L^3$	$\left[\frac{\sin k_1(a/2\sqrt{2})l_1}{l_1 \sin k_1(a/2\sqrt{2})}\right]^2 \left[\frac{\sin k_2(\sqrt{3}a/2)l_2}{l_2 \sin k_2(\sqrt{3}a/2)}\right]^2 \left[\frac{\sin k_3(\sqrt{3}a/2\sqrt{2})l_3}{l_3 \sin k_3(\sqrt{3}a/2\sqrt{2})}\right]^2$	$\exp(-L^2k^2/4\pi)$	$T = 0$	$0{,}2\ k\theta$	$0{,}16\ k\theta^*$ $=0{,}16\ k\theta(a/L)$	—
			$T \gg \theta$	$0{,}47\ kT$		
			$T \gg \theta^*$ $= \theta(a/L)$		$0{,}47\ kT$	
Quader $V = L^2\sqrt{a/3}$	$\left[\frac{\sin k_1(a/2\sqrt{2})l_1}{l_1 \sin k_1(a/2\sqrt{2}}\right]^2 \left[\frac{\sin k_3(\sqrt{3}a/2\sqrt{2})l_3}{l_3 \sin k_3(\sqrt{3}a/2\sqrt{2})}\right]^2$	$\exp[-L^2(k_1{}^2+k_3{}^2)/4\pi]$	$T = 0$	$0{,}2\ k\theta$	$0{,}22\ k\theta$	$0{,}28\ k\theta$
			$T \gg \theta$	$0{,}47\ kT$	kT	$1{,}4\ kT$
Prisma $V = La^2/\sqrt{6}$	$\left[\frac{\sin k_3(\sqrt{3}a/2\sqrt{2})l_3}{l_3 \sin k_3(\sqrt{3}a/2\sqrt{2})}\right]^2 \cos^2 k_1 a/4\sqrt{2}$	$\frac{1}{2}\exp(-L^2k_3{}^2/4\pi)$	$T = 0$	$0{,}2\ k\theta$	$0{,}23\ k\theta$	—
			$T \gg \theta$	$0{,}47\ kT$	$0{,}5\ kT$	—

VOL. 4 1958

die klassische Formel hier schom für viel tiefere Temperaturen gilt, nämlich für $T \gg \theta^*$ mit $\theta^* = \theta(a/L)$. Alle Ausdrücke lassen sich somit in der Form

$$\overline{\tau^2} = zg\frac{E}{v} \tag{3.8a}$$

schreiben, wo z ein Zahlenfaktor der Grössenordnung eins, v das Probevolumen und E die thermische Energie der Grenzschwingung

$$E = \begin{cases} \dfrac{k\theta}{2} & (T = 0) \\ kT & (T \gg \theta) \end{cases} \tag{3.8b}$$

bedeuten. Hierin ist beim Würfel θ^* anstelle von θ zu verwenden. $\overline{\tau^2}/2g$ ist übrigens die mittlere thermische Energiedichte der mittleren Schubspannung.

Das Wesentliche an diesem Ergebnis besteht darin, dass das Schwankungsquadrat der über ein Probevolumen gemittelten Schuspannung umgekehrt proportional mit diesem Volumen abnimmt, und dass der Proportionalitätsfaktor z von der Grössenordnung eins ist.

Das über den Würfel gemittelte Schwankungsquadrat nimmt insofern eine Sonderstellung ein, als dieser Ausdruck für $T = 0$ stärker mit dem Volumen abnimmt, nämlich mit der 4/3-ten Potenz. Das ist unmittelbar aus dem in Abb. 5 dargestellten Korrelationsverlauf für die Verzerrung $\varepsilon_{12}{}^n$ ersichtlich. Im isotropen Fall ist ja $\varepsilon_{12}{}^n = (1/2g)\tau_{12}{}^n$, sodass Abb. 5 auch den Korrelationskoeffizienten der Schubspannung $\tau_{12}{}^n$ innerhalb der x_1, x_2— Ebene und senkrecht dazu angibt.† Wie oben schon bemerkt, ist der Korrelationsverlauf senkrecht zur x_1, x_2—Ebene praktisch temperaturunabhängig, während innerhalb dieser Ebene bei $T = 0$ die negative Korrelation stärker als bei hoher Temperatur vertreten ist. Die Werte des Schwankungsquadrates für Quader und Würfel unterscheiden sich aber gerade darin, dass beim Quader die Korrelation senkrecht zur x_1, x_2—Ebene und längs einer Strecke L innerhalb dieser Ebene eingeht, während beim Würfel die Korrelation senkrecht zur x_1, x_2—Ebene und auf einer ganzen Fläche L^2 innerhalb dieser Ebene eingeht. Der Wert des Würfels ist also stärker von dem temperaturabhängigen Teil der Korrelation bestimmt und fällt also bei $T = 0$ wegen deren grösserem negativen Anteil schneller mit dem Volumen ab als bei hoher Temperatur.

† Dass zur Berechnung der in Abb. 5 dargestellten Korrelationsfunktion eine andere Lage des Koordinatensystems bezüglich der Elementarzelle als in diesem Abschnitt vorausgesetzt wurde, hat im Fall der Isotropie keinen Einfluss auf das Resultat.

In Spalte 5 von Tabelle 4 sind zum Vergleich die ohne Korrelation berechneten Werte (3.2) des Schwankungsquadrates eingetragen. Der in Spalte 7 angegebene Wert wurde durch angenäherte Berücksichtigung des richtigen gittertheoretischen Dispersionsverlaufes im isotropen Fall berechnet.(7)

(c) *Diskussion*

Die Schwankungsquadrate der Spalten 5 und 6 stimmen für $T \gg \theta$ im wesentlichen überein. Die Korrelation hat also fast keinen Einfluss auf das Ergebnis. Das bedeutet, dass man allgemein bei hohen Temperaturen die Korrelation vernachlässigen darf, wenigstens soweit Volumenmittelwerte interessieren. Für $T = 0$ stimmen die mit und ohne Korrelation berechneten Schwankungsquadrate nur beim Prisma und Quader grössenordnungsmässig überein; der mit Korrelation berechnete Wert ist dagegen beim Würfel um einen Faktor (a/L) kleiner als der bei statistischer Unabhängigkeit gefundene. Die Wirkung der Korrelation ist hier also sogar negativ. Allgemein wird man sagen dürfen, dass die Korrelation für tiefe Temperaturen bei makroskopischen Volumina berücksichtigt werden muss, aber bei Volumina, die in mindestens einer Dimension nur atomare Ausdehnung besitzen, vernachlässigt werden darf.

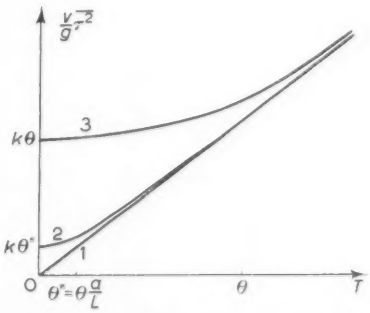


ABB. 7. Temperaturverlauf des Schwankungsqu adrates der Verzerrung eines quaderförmigen Probevol umens.
(1). Beliebiges Probevolumen mit Stempel; klassisch
(2). Würfel ohne Stempel; quantentheoretisch
(3). Flacher Quader und Prisma ohne Stempel; quantentheoretisch.

In Abb. 7 ist der ungefähre Temperaturverlauf der Schwankungsquadrate dargestellt, wie er sich durch Interpolation zwischen dem Wert des Schwankungsquadrates bei $T = 0$ und bem bei hohen Temperaturen gültigen linearen Anstieg ergibt.

Alle Zahlenfaktoren der Grössenordnung eins wurden willkürlich gleich eins gesetzt. Für Quader und Prisma fallen die mit und ohne Korrelation berechneten Kurven (im wesentlichen) zusammen (Kurve 3). Alle als Kurve 3 zusammengefassten Kurven gehen etwa bei der Debye-Temperatur θ in den linearen Verlauf über, welcher der thermodynamischen Formel (1.1b) entspricht (Kurve 1). Die unter Berücksichtigung der Korrelation für den Würfel gültige Kurve weicht jedoch erst bei der tieferen Temperatur $\theta^* = \theta(a/L)$ vom linearen Verlauf ab. Die Korrelation erweitert also die Gültigkeit des linearen kT—Gesetzes beim Würfel, aber nicht beim Quader und Prisma.

Die Werte der Spalte 6 von Tabelle 4 unterscheiden sich von dem Wert (1.1b) der Beckerschen Formel einmal durch die Zahlenfaktoren, sodann durch die andere Volumenabhängigkeit beim Würfel für $T = 0$ und schliesslich durch die andere Temperaturabhängigkeit bzw. das Auftreten eines endlichen Schwankungsquadrates für $T = 0$.

Was die Zahlenfaktoren anbelangt, so ist eigentlich zu erwarten, dass beim Würfel für hohe Temperatur das mit Korrelation berrechnete Schwankungsquadrat in Spalte 6 genau mit dem Wert (1.1b) der Beckerschen Formel übereinstimmt. Der Zahlenfaktor 0,47 ist also offenbar eine Folge der vereinfachenden Annahme, die erst die Berechnung des Schwankungsquadrates ermöglichten. Und zwar ist es gerade die Ersetzung der strengen Volumenfaktoren in Spalte 2 durch blosse Exponentialfunktionen, die einen zu kleinen Wert für das Schwankungsquadrat liefert.

Die Berücksichtigung der Korrelation liefert beim Würfel für hohe Temperatur den gleichen Wert 0,47 wie ohne Korrelation; diese Übereinstimmung ist also rein zufällig.

Aus dem in Spalte 7 angegebenen Wert ersieht man, dass die Verwendung des richtigen Spektrums den Zahlenfaktor nur wenig vergrössert, aber nicht zu einer anderen Volumenabhängigkeit führt. Dasselbe gilt auch von der Berücksichtigung der Anisotropie.[7]

Da die Verwendung der strengen Volumenfaktoren in (3.7) beim Würfel einen um einen Faktor zwei grösseren Wert des Schwankungsquadrates liefern würde, sind offenbar die anderen Schwankungsquadrate für den Quader und das Prisma in Spalte 6 ebenfalls nur bis auf einen solchen Zahlenfaktor richtig. Für $T \gg \theta$ besitzen aber alle diese Werte die gleiche Grössenordnung wie der Wert (1.1b) der Beckerschen Formel; ausserdem hängen diese statistisch berechneten Schwankungsquadrate nicht von der Art des Probevolumens, sondern nur von dessen Grösse ab, und zwar in derselben Weise wie der Wert (1.1b). Die statistische Begründung führt also im wesentlichen (für $T \gg \theta$) in allen Fällen wieder auf die Beckersche Formel.

Das bedeutet, dass die Korrelation für die Beckersche Formel keinen günstigeren Ausdruck liefert, der etwa einen grösseren Wert für das kritische Volumen bei der Ableitung ergibt. Lediglich die Temperaturabhängigkeit der Beckerschen Formel wird durch die Korrelation geändert.

Das statistisch berechnete Schwankungsquadrat stimmt nur für $T \gg \theta$ mit dem linearen Temperaturverlauf (1.1b) überein. Für $T \leqslant \theta$ bezw. $T \leqslant \theta^*$ treten Abweichungen auf (Abb. 7). Hier erhebt sich die Frage nach dem Ursprung dieser Abweichungen. Einerseits ist nämlich die statistische Formel von voller Allgemeingültigkeit. Andererseits ist doch aber auch zu erwarten, dass die thermodynamisch begründete Formel (1.1) unbegrenzte Gültigkeit besitzt, weil ja die Thermodynamik für hinreichend grosse Systeme—also etwa für grosse würfelförmige Probevolumen—bei allen Temperaturen gilt.

Die Lösung dieses Widerspruches besteht in der Einsicht, dass die statistisch und die thermodynamisch begründete Formel zwei verschiedene physikalische Sachverhalte beschreiben, die nur im klassischen Grenzfall hoher Temperatur übereinstimmen. Bei der Ableitung der thermodynamischen Formel (1.1) wurde nämlich vorausgesetzt, dass die Scherung ϵ des Probevolumens makroskopisch fixiert werden kann, dass also ϵ als Parameter betrachtet werden darf. Das ist aber nur dann möglich, wenn die Gestalt des Probevolumens durch einen starren makroskopischen „Stempel" festgelegt wird, dessen Masse gross gegenüber der Masse der Gitteratome ist. Im vorliegenden Fall würde dieser „Stempel" aus vier

starren Blechen bestehen, die in das Kristallgitter eingebettet und deren Kanten derart gelenkig miteinander verbunden sind, dass dadurch als einzige Bewegung unter dem Einfluss der thermisch angeregten Gitterschwingungen die Scherung dieser Anordnung im ganzen möglich ist. Die vier schraffierten Rechtecke in Abb. 1 können unmittelbar als diese materiellen Bleche, die vier Kanten, in denen sie sich berühren, als Gelenke aufgefasst werden.

Die Verhältnisse lassen sich am Modell einer freien linearen Kette besonders einfach darstellen und rechnen. Will man die thermischen Schwankungen der Kettenlänge behandeln, so hat man die beiden Möglichkeiten, an den Enden der Kette zwei grosse Massen als Stempel anzubringen, die gross gegenüber den Massen der Kettenatome sind, oder aber eine gewöhnliche lineare Kette zu verwenden, deren Randatome die gleiche Masse wie die inneren Atome haben. In ersten Fall ist die Kettenlänge wegen der grossen Endmassen ein makroskopisch fixierbarer Parameter, für dessen Schwankungsquadrat man über die freie Energie das lineare kT-Gesetz erhält (Abb. 7, Kurve 1). Der zweite Fall liefert dagegen den in Kurve 3 dargestellten Temperaturverlauf; die Kette besitzt dann auch Nullpunktsschwankungen.

Tatsächlich muss auch die Bewegung des Stempels bei sehr tiefer Temperatur quantentheoretisch behandelt werden, sodass auch hier Abweichungen vom linearem Temperaturverlauf auftreten. Durch Wahl von hinreichend grossen Endmassen der Kette lässt sich jedoch die entsprechende Debye-Temperatur des Stempels beliebig klein machen, sodass in diesem Fall für alle messtechnisch überhaupt zugänlichen Temperaturen das lineare kT-Gesetz gilt.

Aus den obigen Überlegungen ist ersichtlich, dass nur die statistische, nicht die thermodynamische Formel für die Verteilung der mittleren Schubspannung allen Betrachtungen im Kristall zugrundegelegt werden muss. Denn im wirklichen Kristall ist kein Stempel vorhanden; wirklich vorhanden sind dagegen die mittlere Verzerrung des Probevolumens und die dazu proportionale mittlere Schubspannung. Allein durch diese Grössen werden alle dynamischen Vorgänge, die etwa bei der Erzeugung oder Bewegung von Versetzungen eine Rolle spielen, bestimmt.

(d) *Ausblick*

Die Bedeutung der obigen Überlegungen liegt einmal in dem Nachweis, dass die Statistik die Beckersche Formel für hohe Temperaturen bestätigt. Damit ist die Vermutung, dass eine bessere Begründung die Beckersche Formel modifizieren und bei ihrer Anwendung auf die plastischen Erscheinungen zu einem grösseren Wert für das kritische Probevolumen führen würde, widerlegt.

Weiter ist die Beckersche Formel nun für alle Temperaturen statistisch begründet worden. Ausserdem ist gezeigt, dass für die Statistik bei hohen Temperaturen die Korrelation keine Rolle spielt. Auch bei tiefen Temperaturen darf die Korrelation vernachlässigt werden, sofern Volumina interessieren, die in mindestens einer Dimension nur atomare Ausdehnung haben. In diesen Fällen ist etwa die Verteilung der Schubspannungen durch ein Produkt einfacher Gauss-Verteilungen mit dem Schwankungsquadrat (2.30) darstellbar. Lediglich für makroskopische Volumina muss bei tiefen Temperaturen die Korrelation, die hier einen negativen Effekt zeigt, berücksichtigt werden. Das bedeutet, dass für die Verteilung der Schubspannungen die allgemein gültige Formel (2.26) zu verwenden ist.

Diese Bemerkungen gelten zunächst nur für die Schubspannungen. Bei den Verschiebungen ist dagegen wegen des langsamen Abfalls der Korrelation mit der Entfernung (Abb. 4) zu erwarten, dass in jedem Fall die Korrelation zu berücksichtigen, also die strenge Verteilung (2.9) zu verwenden ist.

Eine weitere Folgerung aus den obigen Überlegungen lässt sich an die Grösse des statistisch berechneten Schwankungsquadrates anknüpfen. Während das thermodynamisch berechnete Schwankungsquadrat und das statistisch berechnete Schwankungsquadrat für einen makroskopischen Würfel am absoluten Nullpunkt verschwindet bzw. grössenordnungsmässig kleiner als im klassischen Fall wird, haben die statistisch berechneten Schwankungsquadrate für Probevolumina, die in mindestens einer Dimension nur atomare Ausdehnung besitzen, die gleiche Grössenordnung wie im klassischen Fall.* Die Nullpunktsschwingungen werden also dynamisch in der gleichen

* Hierauf hat bereits CRUSSARD hingewiesen.[8]

Weise und fast im gleichen Masse auf eine Versetzung wirken können, wie die gewöhnlichen thermischen Schwingungen. Diese Wirkung besteht darin, dass eine Versetzung im thermischen Gleichgewicht mit dem Kristall steht, und dass es Vorgänge wie die Diffusion einer Versetzungslinie oder einer Quelle über ein Hindernis hinweg gibt.

Somit ist das bei $T = 0$ auftretende endliche Schwankungsquadrat auch eine Warnung, die bekannten thermodynamischen Schwankungsformeln leichtfertig auf solche Vorgänge bei tiefer Temperatur anzuwenden, weil eben hierbei in der Thermodynamik die Existenz eines „Stempels" bzw. die makroskopische Fixierbarkeit des entsprechenden Parameters vorausgesetzt wird.*

4. ANHANG

(a) *Berechnung der Korrelationsfunktionen aus Abschnitt* 2

Die Korrelationsfunktion der Verschiebungen erhält man durch Einsetzen von (2.11) in (2.10a.) Setzt man darin entsprechend der Debyeschen Näherung $\omega_{k\lambda} = c_\lambda k$ und $e_{1i}{}^{k} = k_i/k$ ein, wobei c_λ wie oben die Schallgeschwindigkeitenbezeichnet, so erhält man unter Verwendung von (2.2b) für die beiden Grenzfälle $T = 0$ ($\epsilon_{k\lambda} = \frac{1}{2}\hbar c_\lambda k$) und $kT \gg \hbar\omega_g$ ($\epsilon_{k\lambda} = kT$) :

$$s_{ik}{}^{k} = \begin{cases} \dfrac{\hbar}{2mc_2}\left\{(\delta-1)\dfrac{k_i k_k}{k^2}+\dfrac{1}{k}\delta_{ik}\right\} & (T=0) \\ \dfrac{kT}{mc_2{}^2}\left\{(\delta^2-1)\dfrac{k_i k_k}{k^4}+\dfrac{1}{k^2}\delta_{ik}\right\} & (kT \gg \hbar\omega_g) \end{cases} \tag{4.1}$$

δ ist das Verhältnis der Schallgeschwindigkeiten, für das wieder der Wert 1/2 verwandt werden soll.

Zur Berechnung von (2.10a) wird die Summation durch eine Integration über $\boldsymbol{k}$ ersetzt.

Zur Ausführung der Integration werden im k-Raum Polarkoordinaten eingeführt und die 1. Brillouin-Zone durch eine Kugel vom Radius k_g (mit $(4\pi/3)k_g{}^3 = v_{\mathrm{Br}}$) ersetzt.

Wir interessieren uns speziell nur für $S_{11}(\boldsymbol{\nu})$. Legt man die Polarachse in die Richtung von $\boldsymbol{\nu}$, so lassen sich die drei Integrationen leicht ausführen und man erhält :

$$S_{11}(\boldsymbol{\nu}) = \begin{cases} \dfrac{\hbar}{2mc_2}\dfrac{2\pi}{v_{\mathrm{Br}}}\dfrac{1}{r^2}\left\{\left(1+2\dfrac{x_1{}^2}{r^2}\right)-\left(2-\dfrac{x_1{}^2}{r^2}\right)\cos\tau+\left(1-3\dfrac{x_1{}^2}{r^2}\right)\dfrac{\sin\tau}{\tau}\right\} & (T=0) \\ \dfrac{kT}{mc_2{}^2}\dfrac{\pi}{v_{\mathrm{Br}}}\dfrac{1}{r}\left\{\left(5+3\dfrac{x_1{}^2}{r^2}\right)\mathrm{Si}(\tau)+3\left(1-3\dfrac{x_1{}^2}{r^2}\right)\left(\dfrac{\sin\tau}{\tau^2}-\dfrac{\cos\tau}{\tau}\right)\right\} & (kT \gg \hbar\omega_g) \end{cases} \tag{4.2}$$

x_i sind die Komponenten von $\boldsymbol{\nu}$ in dem Koordinatensystem der Abb. 2; weiter ist $r = |\boldsymbol{\nu}|$ und $\tau = k_g\, r$.

Si(τ) bezeichnet den Integralsinus :

$$\mathrm{Si}(\tau) = \int_0^\tau \frac{\sin\alpha}{\alpha}\, d\alpha.$$

Für $\boldsymbol{\nu} = 0$ erhält man daraus die Schwankungsquadrate :

$$S_{11}(0) = \begin{cases} \dfrac{5}{48}\left(\dfrac{3}{\pi}\right)^{2/3}\dfrac{\hbar a}{mc_2} & (T=0) \\ \dfrac{3}{16\pi}\left(\dfrac{3}{\pi}\right)^{1/3} a^2\dfrac{kT}{mc_2{}^2}. & (kT \gg \hbar\omega_g) \end{cases} \tag{4.3}$$

Damit erhält man für den Korrelationskoeffizienten

$$\gamma_{11}(\boldsymbol{\nu}) = \begin{cases} \dfrac{6}{5}\left\{\left(1+2\dfrac{x_1{}^2}{r^2}\right)\dfrac{1}{\tau^2}-\left(2-\dfrac{x_1{}^2}{r^2}\right)\dfrac{\cos\tau}{\tau^2}+\left(1-3\dfrac{x_1{}^2}{r^2}\right)\dfrac{\sin\tau}{\tau^3}\right\} & (T=0) \\ \dfrac{1}{6}\left\{\left(5+3\dfrac{x_1{}^2}{r^2}\right)\dfrac{\mathrm{Si}(\tau)}{\tau}+3\left(1-3\dfrac{x_1{}^2}{r^2}\right)\left(\dfrac{\sin\tau}{\tau^3}-\dfrac{\cos\tau}{\tau^2}\right)\right\}. & (kT \gg \hbar\omega_g) \end{cases} \tag{4.4}$$

* Herrn Prof. Dr. G. LEIBFRIED habe ich für die Anregung zu dieser Arbeit sowie für zahlreiche Diskussionen zu danken. Darüberhinaus bin ich ihm auch für manche fruchtbare Idee, wie etwa die von der Rolle des Stempels bei der Beckerschen Formel, sehr verbunden.

Gl. (4.4) enthhält in der Debyeschen Näherung sämtliche Korrelationskoeffizienten der Verschiebungen; es ist nämlich $\gamma_{ik}(\boldsymbol{\nu}) = \gamma_{ii}(\boldsymbol{\nu})\delta_{ik}$ und weiter $\gamma_{11}(x_1, x_2, x_3) = \gamma_{22}(x_2, x_3, x_1) = \gamma_{33}(x_3, x_1, x_2)$.

In Abb. 4 sind die beidem Funktionen (4.4) jeweils für die beiden Spezialfälle $\boldsymbol{\nu} = (x_1, 0, 0)$ und $\boldsymbol{\nu} = (0, x_2, 0)$ mit $\tau_i = k_g x_i = 6{,}22(x_i/a)$ als Funktion von x_i/a aufgetragen. Die gestrichelte Vertikale bezeichnet die Lage der nächsten Nachbarn, die sich dann ergäbe, wenn man die x_1-Achse nicht in die Kante, sondern in die Flächendiagonale der Elementarzelle legte. Die entsprechende Abzisse beträgt $x_1/a = 1/\sqrt{2}$.

Aus Abb. 4 erhält man auch unmittelbar die in Tabelle 2 angegebenen Zahlenwerte des Korrelationskoeffizienten; die Zahlenwerte aus Tabelle 3 erhält man, wenn $\gamma_{11}(a, 0, 0) = \lambda$, $\gamma_{11}(0, a, 0) = \mu$, $\gamma_{11}(a/\sqrt{2}, 0, 0) = 2\alpha - \beta$, $\gamma_{11}(0, a/\sqrt{2}, 0) = \beta$ und $\gamma_{ik} = \gamma_{ii}\delta_{ik}$ beachtet wird.

Hat man statt des kubisch flächenzentrierten Gitters ein anderes primitives kubisches Gitter vorzuliegen, so bleibt (4.4) gültig, wenn nur jetzt $\tau = \sqrt[3]{6\pi^2 z}\; r/a$ bedeutet und z die Zahl der Atome im Elementarwürfel angibt ($z = 1, 2, 4$ für primitives, raumzentriertes, flächenzentriertes Gitter). In (4.2) und (4.3) tritt dann ausserdem der Faktor $(4/z)^{1/3}$ für $T = 0$ und $(4/z)^{2/3}$ für $kT \gg \hbar\omega_g$ hinzu.

Für die Korrelationsmatrix der Schubverzerrung gilt nach (2.20a) und (2.22a) :

$$E_{12}(\boldsymbol{\nu}) = \frac{1}{4v_{\mathrm{Br}}} \iiint \{k_1^2 s_{22\boldsymbol{k}} + k_2^2 s_{11\boldsymbol{k}} + 2k_1 k_2 s_{12\boldsymbol{k}}\}\, d\boldsymbol{k}. \tag{4.5}$$

Dividiert man die vier Funktionen (4.6) und (4.7) durch die zugehörigen Schwankungsquadrate

$$E_{12}(0) = \begin{cases} \dfrac{\hbar k_g}{20mc_2} = 0{,}31\dfrac{\hbar}{amc_2} & (T = 0) \\ 0{,}116\dfrac{kT}{mc_2^2} & (kT \gg \hbar\omega_g) \end{cases} \tag{4.8}$$

so erhält man die entsprechenden Korrelationskoeffizienten $e_{12}(x_1, 0, 0)$ und $e_{12}(0, 0, x_3)$, deren Verlauf in Abb. 5 dargestellt ist.

Da im isotropen Fall die Schubspannung τ_{12} der Verzerrung proportional ist, stellt Abb. 5 unmittelbar auch den Korrelationsverlauf hierfür dar.

(b) *Berechnung des Volumenfaktors D_k^2*

Es soll der in Gl. (3.4) definierte Volumenfaktor durch Ausführung der Summe über die in einem Prisma, einem Quader und einem Würfel gelegenen Atome berechnet werden.

Prisma. In Abb. 8(a) ist eine [111] -Ebene des kubisch flächenzentrierten Gitters dargestellt. Die x_1- und x_3- Achse fallen mit dem in der Fussnote auf Seite 40 angegebenen Basissystem zusammen. Eine solche in x_3-Richtung ($(\bar{1}\bar{1}2)$-Richtung)

Setzt man hierin (4.1) ein und führt wieder Polarkoordinaten ein, so lassen sich die Integrationen leicht ausführen und ergeben für die beiden Spezialfälle $\boldsymbol{\nu} = (x_1, 0, 0)$ und $\boldsymbol{\nu} = (0, 0, x_3)$ mit $\tau = k_g . x_i$:

$$E_{12}(x_1, 0, 0) = \begin{cases} \dfrac{3\hbar k_g}{8mc_2}\left\{-\dfrac{\cos\tau}{\tau^2} + 5\dfrac{\sin\tau}{\tau^3} + 16\dfrac{\cos\tau}{\tau^4} + \dfrac{8}{\tau^4} - 24\dfrac{\sin\tau}{\tau^5}\right\} & (T = 0) \\ \dfrac{3kT}{4mc_2^2}\left\{-\dfrac{\cos\tau}{\tau^2} + 5\dfrac{\sin\tau}{\tau^3} + 2\dfrac{\mathrm{Si}(\tau)}{\tau^3} - 18\left(\dfrac{\sin\tau}{\tau^5} - \dfrac{\cos\tau}{\tau^4}\right)\right\} & (kT \gg \hbar\omega_g) \end{cases} \tag{4.6}$$

$$E_{12}(0, 0, x_3) = \begin{cases} \dfrac{3\hbar k_g}{4mc_2}\left\{-\dfrac{\sin\tau}{\tau^3} - 3\dfrac{\cos\tau}{\tau^4} + 3\dfrac{\sin\tau}{\tau^5}\right\} & (T = 0) \\ \dfrac{3kT}{8mc_2^2}\left\{\dfrac{\mathrm{Si}(\tau)}{\tau^3} - 4\dfrac{\sin\tau}{\tau^3} - g\left(\dfrac{\cos\tau}{\tau^4} - \dfrac{\sin\tau}{\tau^5}\right)\right\} & (kT \gg \hbar\omega_g) \end{cases} \tag{4.7}$$

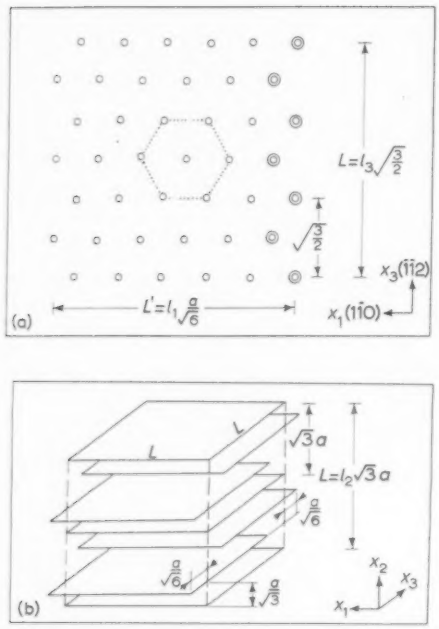


ABB. 8. Prisma, Quader und Würfel, über die bei der Berechnung der Volumenfaktoren D_k^2 summiert wird

(a) [111] -Ebene des kubisch flächenzentrierten Gitters

(b) Die Aufeinanderfolge der [111] -Ebenen im kubisch flächenzentrierten Gitter.

gelegene Kette von Atomen, wie sie in Abb. 8 durch die grossen Kreise angedeutet ist, erfüllt ein prismatisches Volumen, dessen Länge $L=l_3\sqrt{(3/2)}\,a$ und dessen Querschnitt $F=(a/\sqrt{2})\times\times(a/\sqrt{3})=a^2/\sqrt{6}$ beträgt. $a/\sqrt{3}$ ist der Abstand der Netzebenen in x_2-Richtung. Die Zahl der in diesem Prisma enthaltenen Atome ist $n=2l_3$.

Die Summation lässt sich leicht ausführen und ergibt, wenn die k_i auf das angegebene Koordinatensystem bezogen werden :

$$D_k^2=\left[\frac{\sin k_3(\sqrt{3}a|2\sqrt{2})l_3}{l_3\sin k_3(\sqrt{3}a|2\sqrt{2})}\right]^2\cos^2(k_1+\sqrt{3}k_3)\frac{a}{4\sqrt{2}}. \tag{4.9}$$

In dem später benötigten Intervall $|k_3|\sqrt{3}a/2\sqrt{2}<\pi/2$ ist diese Funktion nur für $|k_3|\sqrt{3}a/2\sqrt{2}<\pi/l_3$ wesentlich von null verschieden, sodass für den zweiten Faktor der in diesem Bereich gültige Ausdruck verwendet werden darf. Für $l_3\gg 1$ erhält man somit durch Entwicklung des $\cos^2$-Terms in (4.3) :

$$\cos^2(k_1+\sqrt{3}k_3)\frac{a}{4\sqrt{2}}=\cos^2k_1\frac{a}{4\sqrt{2}}+O(l_3^{-1}) \tag{4.10}$$

Einsetzen in (4.9) liefert den in Spalte 2 von Tabelle 4 angegebenen Volumenfaktor des Prismas.

Quader. Die Summation ist über l_1 Prismen aus 1) zu erstrecken, die zusammen einen Quader mit den Kantenlängen $L=l_3\sqrt{3}/2a$, $L'=l_1a/\sqrt{2}$ und mit der Höhe $a/\sqrt{3}$ bilden (Abb. 8a). Für l_1, $l_3\gg 1$ ist auch bei ganzzahligem l_1, l_3 in guter Näherung $L=L'$, sodass die Basis des Prismas quadratisch wird. Die Summation liefert

$$D_k^2=\left[\frac{\sin k_1(a|2\sqrt{2})l_1}{l_1\sin k_1(a|2\sqrt{2})}\right]^2\times$$

$$\times\left[\frac{\sin k_3(\sqrt{3}a|2\sqrt{2})l_3}{l_3\sin k_3(\sqrt{3}a|2\sqrt{2})}\right]^2\times$$

$$\cos^2(k_1+\sqrt{3}k_3)\frac{a}{4\sqrt{2}}. \tag{4.11}$$

Dieser Ausdruck ist wegen der beiden ersten Faktoren nur für

$$|k_1|\frac{a}{2\sqrt{2}}<\frac{\pi}{l_1}\ll 1,\quad |k_3|\frac{\sqrt{3}a}{2\sqrt{2}}<\frac{\pi}{l_3}\ll 1$$

wesentlich von null verschieden; in diesem Bereich ist jedoch :

$$\cos^2(k_1+\sqrt{3}k_3)\frac{a}{4\sqrt{2}}=1+O(l_1^{-2})O(l_3^{-2}) \tag{4.12}$$

(4.11) ist zusammen mit (4.12) der in Spalte 2 von Tabelle 4 augegebene Volumenfaktor.

Würfel. Der Würfel soll aus $3l$, aufeinanderliegenden Quadern, also aus $3l_2$ im Abstand $a/\sqrt{3}$ aufeinanderfolgenden quadratischen Netzebenenabschnitten bestehen. Aus Abb. 8(b) ist ersichtlich, dass diese Netzebenen gegeneinander in x_3-Richtung um $\pm a/\sqrt{6}$ verschoben sind. Das

Volumen des Würfels beträgt $V = L^3 = l_1 a/(\sqrt{2}) \times$ $\times 3l_2 a/\sqrt{2} l_3 \sqrt{3}/2a = na/4^3$, wenn $n = 6l_1, l_2, l_3$ die Anzahl der darin enthaltenen Atome ist. Die Summation führt auf

$$D_k^2 = \frac{1}{9}\left[\frac{\sin k_1(a/2\sqrt{2})l_1}{l_1 \sin k_1(a/2\sqrt{2})}\right]^2 \times$$
$$\times \left[\frac{\sin k_2(\sqrt{3}a/2)l_2}{l_2 \sin k_2(\sqrt{3}a/2)}\right]^2 \times$$
$$\times \left[\frac{\sin k_3(\sqrt{3}a/2\sqrt{2})l_3}{l_3 \sin k_3(\sqrt{3}a/2\sqrt{2})}\right]^2 \times$$
$$\times \cos^2(k_1 + \sqrt{3}k_3)\frac{a}{4\sqrt{2}} \cdot$$
$$\times \left\{4\left[\cos^2\left(k_2 - \frac{1}{\sqrt{2}}k_3\right)\frac{a}{2\sqrt{3}} + \right.\right.$$
$$+ \cos^2\left(2k_2 + \frac{1}{\sqrt{2}}k_3\right)\frac{a}{2\sqrt{3}} +$$
$$\left.\left. + \cos^2(k_2 + \sqrt{2}k_3)\frac{a}{2\sqrt{3}}\right] - 3\right\}. \qquad (4.13)$$

Für $l_1, l_2, l_3 \gg 1$ besitzen wieder sämtliche $\cos^2$-Terme in dem Intervall, in dem die drei ersten Faktoren wesentlich von null verschieden sind, bis auf Glieder von der Grössenordnung l_i^{-2} den Wert 1, sodass sich (4.13) auf den in Spalte 2 von Tabelle 4 angegebenen Ausdruck reduziert.

In den obigen Formeln für D_k^2 stammen die Faktoren von der Gestalt $(\sin lx/l \sin x)^2$ jeweils von der Summation über das primitive Gitter, während die $\cos^2$-Terme Strukturfaktoren sind. Es ist interessant, dass—ausser beim Prisma—der Einfluss der speziellen Gitterstruktur für grosse Probevolumina ein Effekt höherer Ordnung ist und daher fortgelassen werden kann.

(c) *Zur Berechnung des Schwankungsquadrates* (3.7)

Zu berechnen sind die sechs in Spalte 6 von Tabelle 4 angegebenen Werte von $\overline{\tau^2}$, die man erhält, wenn in (3.7) die beiden Werte für τ^k nach (3.6) und die drei Werte für D_k^2 aus Spalte 3 von Tabelle 4 eingesetzt werden.

Beim Würfel lässt sich die Integration durch Einführung von Polarkoordinaten leicht ausführen, wobei die Integration über k wegen des raschen Abfalls der Exponentialfunktion bis unendlich erstreckt werden darf.

Beim Quader müssen Zylinderkoordinaten eingeführt und die 1. Brillouin-Zone durch einen volumengleichen Zylinder ersetzt werden. Die Höhe dieses Zylinders muss gleich der Ausdehnung der 1. Brillouin-Zone in der Raumdiagonalen, nämlich gleich $2\sqrt{3}(\pi/a)$ sein. Die Integration über den Zylinderradius darf dann wegen der Exponentialfunktion bis unendlich geführt werden.

Beim Prisma werden ebenfalls Zylinderkoordinaten eingeführt und die 1. Brillouin-Zone durch einen volumengleichen Zylinder ersetzt, dessen Höhe gleich der Ausdehnung in der k_3-Achse ($(\bar{1}\bar{1}2)$-Richtung), also etwa gleich $2\pi/a$ ist. Aus der Volumengleichheit ergibt sich der Zylinderradius zu $4\sqrt{\pi}/a$. Wegen des Exponentialfaktors darf hier die Integration über k_3 bis unendlich erstreckt werden.

LITERATUR

1. Becker R. *Phys. Z.* **26,** 919 (1925).
2. Cottrell A. H. *Dislocations and Plastic Flow in Crystals.* Oxford (1953).
3. Leibfried G. *Handbuch d. Physik* Vol. VII. p. 1. Berlin-Göttingen-Heidelberg (1955).
4. Husimi K. *Proc. Phys. Math. Soc. Japan* **22,** 264 (1940).
5. Curien H. Theses, Paris Nr. 2466/3338 (1952).
6. Uhlenbeck G. E. and Wang M. C. *Rev. Mod. Phys.* **17,** 323 (1945).
7. Barsch G. Dissertation, Göttingen (1955).
8. Crussard C. *L'Etat Solide.* Brüssel (1952).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 50–57.

ZUM ENERGIEUMSATZ BEI RÖNTGENPHOTOCHEMISCHEN PROZESSEN IN KBr-KRISTALLEN

ERNST FELDTKELLER

Zweites Physikalisches Institut der Technischen Hochschule, Stuttgart

(*Received* 30 *April* 1957)

Zusammenfassung—Es wurde untersucht, welcher Anteil der absorbierten Energie bei Röntgenbestrahlung eines KBr-Einkristalls in Wärme ver wandelt wird. Für *reines* KBr beträgt dieser Anteil etwa 90%. Für *hydriertes* KBr liegt die Wärmeausbeute bei Zimmertemperatur zu Anfang der Bestrahlung etwa 3% tiefer als beim reinen Kristall und nähert sich der des reinen KBr im Laufe der Verfärbung. Folgerungen für den Anteil der beteiligten optischen und strahlungslosen Elektronenübergänge werden diskutiert.

Nach demselben Verfahren wurde die entstehende Wärme beim Ubergang *optisch* angeregter F-Zentren in den Grundzustand für KCl gemessen. Die Ergebnisse stehen im Einklang mit der durch BECKER und PICK bestimmten Floureszenzausbeute desselben Prozesses.

Abstract—An investigation was made of the percentage converted into heat, of the absorbed energy in a KBr single crystal irradiated by X-rays. For *pure* KBr this rate amounts to about 90 per cent. For *hydrated* KBr the heat yield at room temperature in the beginning of irradiation is about 3 per cent lower than for the pure crystal, and approaches that of pure KBr during colouration. Deductions for the portion of radiating and radiationless electron transitions involved are discussed.

By the same method the heat produced by transitions of *optically* excited F-centres to the ground state is measured in KCl. The results agree with the fluorescence yield of the same process measured by BECKER and PICK.

1. EINLEITUNG

BEI der Absorption energiereicher Strahlung in Kristallen wird ein Teil der absorbierten Energie photochemisch gespeichert (durch Bildung von Störzentren und durch darin stattfindende Elektronenprozesse), ein Teil durch Lumineszenz wieder ausgestrahlt (optische Elektronenübergänge) und ein dritter Teil im Kristall in Wärme umgesetzt (Stossprozesse, strahlungslose Übergänge und dergl.). Unter der „absorbierten" Strahlungsenergie soll also der weder durchgelassene noch gestreute Teil der eingestrahlten Energie verstanden werden.

In *Metallen* geht fast alle absorbierte Energie in Wärme über. Dort kann wegen der starken Absorption im optischen Bereich keine Lumineszenz auftreten. Auch der durch bleibende Elektronenumlagerungen photochemisch gespeicherte Anteil fällt wegen der grossen Zahl freier Elektronen fort. Nur durch Bausteinumlagerung kann ein wenig Energie photochemisch gespeichert werden (radiation damage).

In *Ionen*kristallen dagegen haben photochemische Energiespeicherung und Lumineszenz einen merklichen Anteil am Energieumsatz der Strahlungsabsorption. Für das Beispiel der Röntgenverfärbung von KBr-Kristallen* ist die Ausbeute der verschiedenen photochemischen Prozesse, ermittelt aus der Änderung des Absorptionsspektrums, weitgehend bekannt und z.B. für die F-Zentrenbildung bei reinem KBr durchweg kleiner als $5 \cdot 10^{-3}$ Zentren/eV.[1] Eine grobe Energieabschätzung erhält man, wenn man als Energiespeicherung wenige eV/Zentrum einsetzt. Man sieht, dass dadurch nur ein kleiner Teil der absorbierten Röntgenenergie photochemisch gespeichert wird.

* Über photochemische Prozesse in Alkalihalogenid-Kristallen allgemein siehe die zusammenfassenden Berichte von POHL[2] und SEITZ.[3]

Das Ziel der vorliegenden Arbeit war es, einen Beitrag zur Aufteilung des übrigen, grössten Anteils der absorbierten Energie zu liefern. Es wurde der in Wärme umgesetzte Bruchteil der vom Kristall absorbierten Röntgenenergie kalorimetrisch gemessen. Dieser Bruchteil wird im folgenden „Wärmeausbeute" γ genannt. Auch der Einfluss definiert eingebauter Störzentren auf die Wärmeausbeute wurde untersucht.

2. DAS METHODISCHE PRINZIP DER MESSUNGEN

Vor der Beschreibung der experimentellen Einzelheiten sei hier zunächst die Messmethode grundsätzlich erläutert. Eine Bestimmung der Wärmeausbeute durch Messung der Kristallerwärmung, der Wärmekapazität des Kristalls und der Röntgenbestrahlungsstärke ist prinzipiell natürlich möglich. Um jedoch die leicht zu wesentlichen systematischen Fehlern führenden Absolutmessungen zu umgehen, wurde hier eine andere Methode angewandt, nämlich ein System von Vergleichsmessungen, das die Wärmeausbeute ohne jede Absoluteichung liefert. Dabei wurde die Erwärmung des KBr-Kristalls verglichen mit der Erwärmung einer Vergleichsprobe, in der die absorbierte Röntgenenergie ganz in Wärme umgesetzt wird, d.h. für die die Wärmeausbeute bei Röntgenbestrahlung eins ist. Als Substanz für diese Vergleichsprobe wurde Blei gewählt, für das die Forderung $\gamma = 1$ hinreichend erfüllt sein dürfte.

Das Verhältnis der Erwärmungen dieser beiden Proben liefert aber noch nicht die Wärmeausbeute des KBr-Kristalls, da die Erwärmung jeder Probe noch von deren Wärmekapazität und Wärmeleitung abhängt und die beiden Proben sich darin beträchtlich unterscheiden. Der hierfür in Rechnung zu setzende Faktor wurde ermittelt, indem bei zwei weiteren Messungen eine Strahlung verwandt wurde, die von beiden Proben mit der Wärmeausbeute eins umgesetzt wird. Dazu wurde die Erwärmung beider Proben bei Bestrahlung mit sichtbarem Licht gemessen und ins Verhältnis gesetzt. Der KBr-Kristall erhielt dazu kolloidal assoziierte Farbzentren und die Bleiprobe eine oberflächliche Russ-Schicht. In besonderen Versuchen (siehe die Kleindruckabschnitte in § 3) wurde nachgewiesen, dass die Verwendung der Russschicht (einer materialfremden Oberflächenschicht) hier keinen Fehler verursacht, sowie dass die kolloidalen Farbzentren die röntgenphotochemischen Prozesse in KBr nicht stören.

Zu jeder Messung wurde das zugehörige Absorptionsvermögen bestimmt (§ 5). Jedes Erwärmungsmass W (genaue Definition folgt in § 4) wurde durch Division mit dem zugehörigen Absorptionsvermögen A auf die tatsächlich absorbierte Strahlungsleistung bezogen. Aus den Erwärmungs- und Absorptionsmessungen ergibt sich die Wärmeausbeute γ nach der Formel

$$\gamma = \frac{W_{r,\,\mathrm{KBr}} \cdot W_{s,\,\mathrm{Pb}}}{W_{r,\,\mathrm{Pb}} \cdot W_{s,\,\mathrm{KBr}}} \cdot \frac{A_{r,\,\mathrm{Pb}} \cdot A_{s,\,\mathrm{KBr}}}{A_{r,\,\mathrm{KBr}} \cdot A_{s,\,\mathrm{Pb}}}.$$

Die Indizes r und s bedeuten Messung bei Röntgen- bzw. sichtbarer Bestrahlung, die Indizes KBr und Pb zeigen das Grundmaterial der betreffenden Probe.

Dadurch, dass in obiger Formel jeder Messwert sowohl durch einen anderen, an derselben Probe gemessenen Wert dividiert wird, als auch durch einen anderen bei derselben Bestrahlungsquelle gemessenen, gehen fast alle systematischen Fehler in den Zähler und Nenner in gleichem Masse ein und fallen dadurch im Endergebnis heraus. Dies war ein besonderer Vorteil des gewählten Verfahrens.

3. DIE EXPERIMENTELLEN EINZELHEITEN

Die *Proben* wurden folgendermassen präpariert : Ein aus der Schmelze gezogener, durch Erhitzen im Kaliumdampf stark additiv verfärbter KBr-Kristall wurde langsam von über 500°C auf Zimmertemperatur heruntergetempert. Für die Messungen wurde davon ein $10 \times 4 \times 0{,}5$ mm³ grosses Plättchen zurechtgespalten.

Durch das Tempern werden die Farbzentren in Kolloide umgewandelt. Von den Kolloiden nimmt man an, dass sie alle Energie des absorbierten Lichtes in Wärme umsetzen. Durch Ausmessen des Absorptionsspektrums (ein Beispiel gibt Abb. 1 wieder) vor und nach Röntgenbestrahlung wurde nachgewiesen, dass die Kolloidbande auch durch die Röntgenbestrahlung nicht verändert wird. Die Röntgenbestrahlung hat lediglich einen Aufbau der F-Bande zur Folge, sodass der kolloidhaltige Kristall in seinem Verhalten bei Röntgenbestrahlung ganz dem „reinen" KBr vergleichbar ist. Man darf also den kolloidhaltigen Kristall auch zur Messung von $W_{r,\,\mathrm{KBr}}$ verwenden und das Ergebnis als für den reinen KBr-Kristall gültig ansehen. Die Verwendung derselben Kristallprobe für die Messung bei

Bestrahlung mit Röntgenlicht und sichtbarem Licht erleichtert eine genaue Messung ganz wesentlich (vergl. § 4 unten). Dabei dürfen allerdings die Messungen mit sichtbarem Licht nicht nach den Röntgenmessungen ausgeführt werden, weil sonst die durch die Röntgenbestrahlung entstandenen F-Zentren die Messung von $W_{s,\,KBr}$ verfälschen würden.

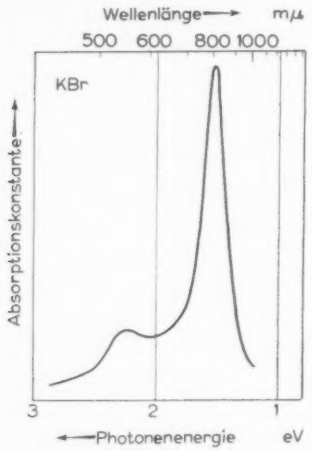


ABB. 1. Absorptionsspektrum eines kolloidal verfärbten KBr-Kristalls.

Ein gleichgrosses Bleiplättchen wurde mechanisch von seiner Oxydschicht befreit und über einer Kampferflamme oberflächlich berusst.

Bei der Messung von $W_{s,\,Pb}$ wurde im Gegensatz zu den anderen Erwärmungsmessungen die zugestrahlte Energie nicht im Innern der Probe selbst, sondern in der aufgebrachten Russ–Schicht absorbiert, über deren Wärmekontakt zum Blei zunächst nichts ausgesagt werden kann. Die Bestimmung der Wärmeausbeute durch die dadurch etwas inkonsequente Verhältnisbildung (siehe § 2) wird durch folgenden Versuch gerechtfertigt: Ein möglicher Fehler müsste sich viel stärker auswirken, wenn man die Messungen statt im Vakuum in Luft unter Atmosphärendruck vornimmt, wo der grösste Teil der gebildeten Wärme direkt von der Probenoberfläche weggeleitet wird, und nicht zum Thermoelement fliesst. Deshalb wurden die Temperaturdifferenzen der thermischen Gleichgewichte bei Röntgen- und sichtbarer Bestrahlung der berussten Bleiprobe in Luft gemessen. Der daraus gebildete Quotient wich nur um 1% von dem im Vakuum gemessenen Quotienten $W_{s,\,Pb}/W_{r,\,Pb}$ ab. Daraus kann geschlossen werden, dass der zunächst befürchtete Fehler die Messungen nicht beeinträchtigt.

Hydrierte KBr-Kristalle wurden aus additiv verfärbtem KBr hergestellt durch Erhitzen unter 30 atm Wasserstoffdruck bis zur Entfärbung. Bei der Bestimmung der Wärmeausbeute des hydrierten KBr wurde an diesem die Grösse $W_{r,\,KBr}$ und an einem kolloidhaltigen KBr-Kristall die Grösse $W_{s,\,KBr}$ gemessen.

Abb. 2 zeigt den *Kristallhalter*, der in einem evakuierbaren Kühltopf untergebracht war. Die Proben waren allseitig von Vakuum umgeben und

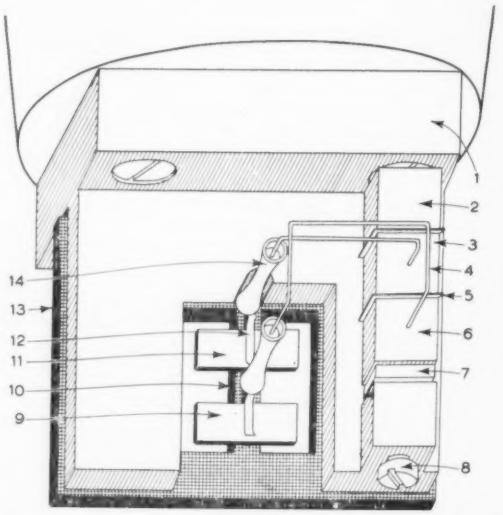


ABB. 2. Vereinfachte Ansicht des Kristallhalters

(1) Kupferner Kristallhalter.
(2) Messingklötzchen, zweite Lötstelle der Konstantanschenkel beider Thermoelemente.
(3) Dasselbe für den Tellurschenkel des oberen Thermoelementes.
(4) Platindraht.
(5) Glimmerisolation.
(6) Wie 3, für unteres Thermoelement.
(7) Pertinax-Zwischenstück, gegen das die Schraube 8 drückt.
(8) Schraube zum Festklemmen der Messingklötzchen.
(9) KBr-Kristallprobe.
(10) Steg der Bleiblende.
(11) Vergleichsprobe aus Blei.
(12) Bronzeklemmfeder, an die das Thermoelement angelötet ist.
(13) Bleiblende.
(14) Tellurgefüllte Glaskapillare.

Die Konstantandrähte der Thermoelemente wurden zur Vereinfachung in der Abbildung weggelassen. Sie führten beide zu demselben Klötzchen 2.

Die Thermoelemente wurden bei der Montage an den Messingklötzchen hängend in den Kristallhalter eingesetzt.

hatten nur über ein Thermoelement die unvermeidliche wärmeleitende Verbindung zum Kristallhalter. Sie hingen in einer Bronzeklemmfeder, an der andererseits das Thermoelement angelötet war. Leitsilber zwischen Probe und Feder sorgte für sicheren Wärmekontakt.

Für die *Thermoelemente* wurde Tellur-Konstantan gewählt wegen der hohen Thermokraft. Das sehr brüchige Tellur war in einer Glaskapillare mechanisch stabil gefasst. Bei der Herstellung der Thermoelemente wurde ein wenig Tellur in einem engen Glasröhrchen geschmolzen und dann das Glas zur Kapillare ausgezogen, wodurch Wärme- und elektrischer Widerstand der Tellursäule auf den gewünschten Wert gebracht wurden. In die über einer kleinen Flamme nochmals geschmolzenen Enden der Tellursäule wurde Platindraht eingeschweisst. An den „zweiten Lötstellen" sorgten $5\times5\times5$ mm³ grosse Messingklötzchen für eine hinreichende Temperaturträgheit. Von ihnen aus führten dünne Kupferleitungen nach aussen zum Galvanometer. Der elektrische Widerstand jedes Thermoelementschenkels mit der Zuleitung betrug 6 bis 7 ohm.

Die Thermoelemente und Bronzefedern wurden durch einen Steg der 2 mm dicken Bleiblende vor direkter Zustrahlung geschützt. Ein Seidenstoffbelag auf dem Steg verhinderte eine grössere Wärmeableitung bei einer zufälligen Berührung zwischen Bronzefeder und Steg.

Die *Röntgenröhre* besass eine W-Antikathode und wurde mit Wechselspannung (90kV Spitze) betrieben. Der mittlere Elektronenstrom der Röhre betrug 7 mA und wurde magnetisch konstant gehalten und, ebenso wie die Primärspannung des Hochspannungstransformators, während der Messungen laufend überwacht. Der weiche Anteil der Bremsstrahlung wurde vor allem in einem Glasfenster des Strahlenschutzmantels absorbiert, sodass die Wellenlänge der Röntgenstrahlung am Messort im Wesentlichen zwischen 0,14 und 0,5 Å gelegen sein dürfte. Als *Lichtquelle* im sichtbaren Spektralgebiet diente eine Glühlampe mit einem Filter, von dem nur der Spektralbereich zwischen 0,7 und 1,3 μ Wellenlänge durchgelassen wurde. Der zeitliche Verlauf der Thermospannungen wurde mit einem *Spiegelgalvanometer* verfolgt. Seine Empfindlichkeit betrug 0,3 μV pro mm/m.

4. AUSWERTUNG DER GEMESSENEN TEMPERATURVERLÄUFE; DEFINITION DER MESSGRÖSSE W

Ein Beispiel für den Verlauf der Thermospannung vor, während und nach der Bestrahlung zeigt Abb. 3. Die Temperatur der Probe nähert sich in jedem Teil der Kurve nach einer Exponentialfunktion einem Gleichgewichtswert (Gleichgewicht zwischen der Wärmeleistung, die durch die Einstrahlung entsteht, und dem Wärmeverlust durch Leitung und Wärmestrahlung). Der Abstand der Gleichgewichtstemperaturen hängt stark von der Wärmeisolierung der Probe ab, d.h. vor allem von der jeweiligen Qualität des Vakuums im Kühltopf. Er repräsentiert deshalb nicht quantitativ die Leistrung der durch Einstrahlung entstehenden Wärme.

Eine besser geeignete Messgrösse wurde folgendermassen gefunden: Der Galvanometerausschlag ϕ wurde numerisch nach der Zeit differenziert und $d\phi/dt$ über ϕ graphisch aufgetragen

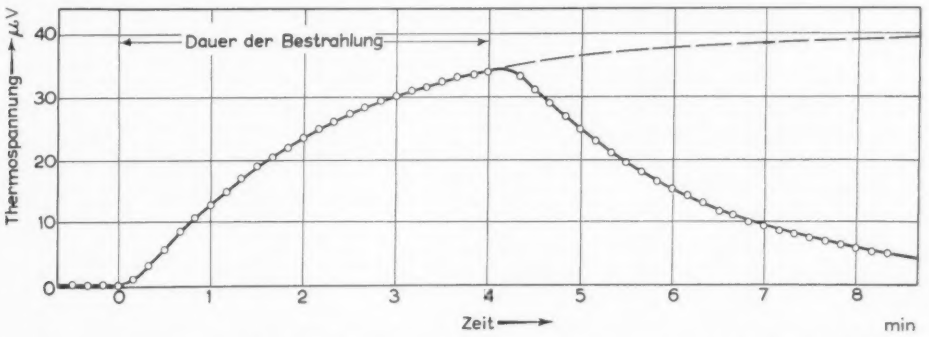


ABB. 3. Zeitlicher Verlauf der Thermospannung vor, während und nach der Bestrahlung.

(Abb. 4). Die Kurven münden nach einem kurzen, durch die Wärmekapazität der Feder usw. bedingten Anfangsverlauf in zwei parallele Geraden ein. Die Geraden kommen dadurch zustande, dass die

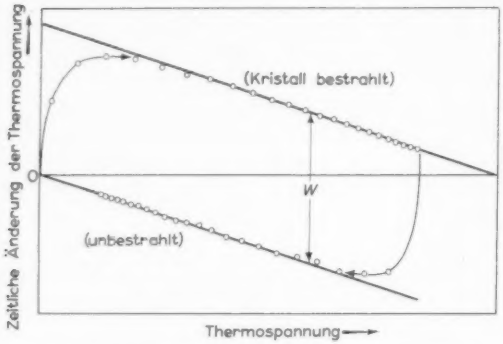


ABB. 4. Graphische Auswertung des in Abb. 3 gezeigten Verlaufs. Näheres siehe Text.

zeitliche Änderung der Temperatur dem momentanen Abstand von der angestrebten Gleichgewichtstemperatur proportional ist.

Der Schnittpunkt der unteren Geraden mit der Abszisse entspricht der Temperaturdifferenz null zwischen Probe und Kristallhalter. Da bei dieser Temperaturdifferenz keine Wärmeleitung von der Probe zum Kristallhalter stattfindet, zeigt der Ordinatenabschnitt der oberen Geraden, wie schnell die Temperatur bei Bestrahlung steigen würde, wenn keine Wärme abgeleitet würde. Als vertikaler Abstand der beiden Geraden an irgend einer Stelle abgelesen, diente er in dieser Arbeit als das in § 2 eingeführte Erwärmungsmass W. Er hat sich tatsächlich als ein reproduzierbares (vom Vakuum unabhängiges) Mass für die Erwärmungsleistung erwiesen. Dafür war allerdings wichtig, dass das temperaturleitende System (vor allem der Kontakt zwischen Feder und Probe) zwischen den Messungen mit Röntgenlicht und mit sichtbarem Licht am gleichen Objekt sich nicht änderte, denn die Wärmeleitverhältnisse längs dieses Systems bestimmten den am Thermoelement abfallenden Teil der gesamten Temperaturdifferenz zwischen Probe und Kristallhalter. Da die zusammengehörenden Messungen zeitlich direkt aufeinander folgten, war dies hinreichend gewährleistet.

5. DIE BESTIMMUNG DER ABSORPTIONSVERMÖGEN

Die Absorptionsvermögen A wurden aus Durchlässigkeit, Reflexion und Streuung bestimmt, wobei die beiden letzteren nur kleinere Korrektionen darstellen.

Für die *Durchlässigkeits*messungen wurde das Messobjekt in den Strahlengang der in § 3 beschriebenen Anordnung gebracht. Als Strahlungsempfänger diente ein geschwärztes Bleiplättchen in der Feder des Thermoelements.

Das *Reflektions*vermögen des KBr-Kristalls bei 800 mμ (der Wellenlänge, bei der das Absorptions-maximum der Kolloide in KBr liegt) wurde aus dem Brechungsindex zu 0,046 berechnet.

Die optische *Streuung* wurde mit einer rotempfindlichen Photozelle gemessen. Sie betrug, über den ganzen Raumwinkel integriert, für das berusste Bleiplättchen 2%, für den kolloidhaltigen KBr-Kristall 5–6% der einfallenden Strahlungsleistung.

Angaben über die Röntgenstreuung wurden, teilweise durch Interpolation, dem Tabellenwerk LANDOLT-BÖRNSTEIN entnommen. Es ergab sich, dass der Anteil des Röntgenstreukoeffizienten am Röntgenschwachungskoeffizienten im hier interessierenden Wellenlängenbereich für KBr etwa derselbe ist wie für Pb. In der Formel von § 2 werden $A_{r,\mathrm{Pb}}$ und $A_{r,\mathrm{KBr}}$ miteinander dividiert. Deshalb konnte bei den Messungen mit Röntgenstrahlen die Streuung überhaupt ausser Betracht gelassen werden.

6. ERGEBNISSE

Aus den Erwärmungs- und Absorptionsmessungen ergab sich für den *reinen* KBr-Kristall eine „Wärmeausbeute" von etwa 90%. Die Werte aus Messungen an verschiedenen Proben streuten zwischen 88,6 und 92,7% mit dem Mittelwert bei 90,6%. Einzelmesswerte sind in Tabelle 1 zusammengefasst. Eine zeitliche Änderung wurde nicht festgestellt. Dies war auch nicht zu erwarten, da die photochemischen Prozesse beim reinen Kristall schon nach kurzer Bestrahlungsdauer einen verschwindend geringen Anteil am Röntgenenergieumsatz haben.

Um die Temperaturabhängigkeit der Wärmeausbeute zu ermitteln, wurden die Quotienten $W_{r,\mathrm{KBr}}/W_{s,\mathrm{KBr}}$ bei Zimmertemperatur und beim Siedepunkt des Sauerstoffs ermittelt und miteinander verglichen. Einzelwerte dazu zeigt Tabelle 2. Durch die Werte W_s wurde die starke Temperaturabhängigkeit der Thermokräfte und der Wärmeleitung berücksichtigt. Verständlicherweise ist die Ungenauigkeit der kalorimetrischen Messungen bei tiefen Temperaturen grösser als bei Zimmertemperatur, sodass in den Ergebnissen lediglich eine

Tabelle 1. *Einzelne Messwerte für die „Wärmeausbeute" der Röntgenphotochemie in KBr-Kristallen bei Zimmertemperatur*

$W_{r,\,KBr}$	$W_{r,\,Pb}$	$W_{s,\,KBr}$	$W_{s,\,Pb}$	$A_{r,\,KBr}$	$A_{r,\,Pb}$	$A_{s,\,KBr}$	$A_{s,\,Pb}$	γ_{+20}
0,154	0,444	0,243	0,535	0,464	0,97	0,56	0,98	0,913
0,170	0,385	0,310	0,490	0,36	0,97	0,46	0,98	0,886
0,182	0,412	0,356	0,490	0,39	0,97	0,58	0,98	0,897
0,148	0,226	0,317	0,258	0,35	1,00	0,595	0,98	0,927

Die ersten vier Spalten enthalten das in § 4 und durch Abb. 4 definierte Erwärmungsmass W für die verschiedenen Messungen, angegeben in μV/sec. Die Spalten 5 bis 8 enthalten das Röntgen-Schwächungs- und das auf das gefilterte Licht bezogene Absorptions-Vermögen. (Wie in § 5 erwähnt, konnte die Röntgenstreuung bei den Messungen unberücksichtigt bleiben). Die letze Spalte enthält die daraus nach der Formel von § 2 errechnete Wärmeausbeute für +20°C.

Jede Zeile gilt für ein miteinander gemessenes Probenpaar.

Die Grösse von $W_{r,\,KBr}$ entspricht etwa einer extrapolierten Anfangs-Temperaturänderung von $1{,}5 \cdot 10^{-3}$ Grad/sec. Im thermischen Gleichgewicht bei Röntgenbestrahlung betrug die Temperaturdifferenz zwischen Kristall und Kristallhalter etwa 0,2 Grad.

Tabelle 2. *Einzelne Messwerte zum Vergleich der Wärmeausbeute bei* −180°C *mit der bei Zimmertemperatur gemessenen*

$W_{r,-180}$	$W_{s,-180}$	$W_{r,+20}$	$W_{s,+20}$	$\dfrac{W_{r,-180} \cdot W_{s,+20}}{W_{s,-180} \cdot W_{r,+20}} = \dfrac{\gamma_{-180}}{\gamma_{+20}}$
0,106	0,15	0,181	0,263	1,03
0,085	0,118	0,166	0,245	1,06
0,094	0,12	0,189	0,238	0,99
0,087	0,116	0,167	0,228	1,02
0,085	0,123	0,177	0,256	1,00

Bedeutung der Buchstaben wie in Tabelle 1. Die Indizes r und s heissen Röntgen- bzw. sichtbare Bestrahlung. Alle Werte sind an KBr gemessen. Die Zahlenindizes geben die Kristalltemperatur in °C an.

Begrenzung des Unterschiedes zwischen der Wärmeausbeute bei Zimmerund bei tiefer Temperatur gesehen werden darf. D.h., ein Anstieg der Wärmeausbeute zu tiefen Temperaturen hin ist nach den Messungen möglich, aber nicht sicher.

Die Messergebnisse für *hydriertes* KBr bei Zimmertemperatur zeigt Abb. 5: Die Wärmeausbeute steigt im Laufe der Verfärbung um etwa 3% an.

Relativ kleine Anderungen in der Wärmeausbeute des hydrierten Kristalls konnten dadurch erfasst werden, dass in den beiden Thermo-elementen ein hydrierter und ein reiner Kristall gleichzeitig bestrahlt wurden. Das Galvanometer wurde während ein und derselben

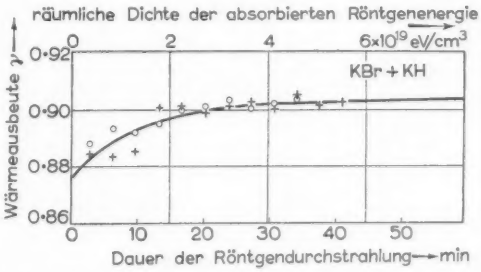


ABB. 5. Verlauf der Wärmeausbeute in hydriertem KBr bei Zimmertemperatur, an zwei verschiedenen Proben gemessen. Anfangs-Hydridkonzentration $1{,}8 \cdot 10^{18}$ U-Zentren/cm³. Durch die Ordinatenteilung sind die Messungen an die Eichung für reines KBr angeschlossen.

Bestrahlungsperiode* nach jeder einzelnen Ablesung auf das andere Thermoelement umgeschaltet. Durch Quotientenbildung $W_{r,\,\mathrm{KBr\ hydr.}}/W_{r,\,\mathrm{KBr\ rein}}$ für jede einzelne Periode wurden kleine Schwankungen in der Strahlungsleistung oder in anderen Einflüssen, deren allmähliche Änderung nicht ganz auszuschliessen ist, eliminiert. Die Wärmeausbeutemessung am hydrierten KBr wurde sozusagen dauernd auf die gleichzeitige Messung am reinen KBr (bei dem γ schon nach sehr kurzer Bestrahlungsdauer zeitlich konstant ist) normiert.

7. DISKUSSION

Für eine eingehendere Betrachtung des röntgenphotochemischen Energieumsatzes ist es zweckmässig, zwei aufeinanderfolgende Prozesse zu unterscheiden:

I. Die Aufteilung der grossen Energie des absorbierten Quants in kleinere Energiebeträge und dadurch Bildung einer grossen Anzahl freier Elektronen und Defektelektronen an der Kante des Leitungs- bzw. Valenzbandes, und

II. den Einfang dieser Elektronen und Defektelektronen in Störstellen, wodurch entweder Rekombination der entgegengesetzten Ladungsträger an der Störstelle stattfindet oder eine Festlegung der Ladungsträger unter Bildung stabiler neuer Störzentrenarten.

Für den Prozess I, die Paarbildung in KBr durch Röntgenbestrahlung, haben MARTIENSSEN und PICK,[4] MIESSNER[5] und LÜTY[1] eine temperaturunabhängige Ausbeute von etwa 1 Paar/40 eV gemessen, obwohl die Breite der verbotenen Zone in KBr nur etwa 8 eV beträgt. Es sind also in den Paaren nur etwa 20% der absorbierten Röntgenenergie potentiell gespeichert. Da man beim Prozess I Verluste durch Strahlung ausschliessen kann, werden die restlichen 80% der absorbierten Röntgenenergie als Wärmeschwingung ans Gitter abgegeben.

Beim Prozess II, der Rekombination der Ladungsträger bzw. deren Festlegung in Störzentren, sind dagegen strahlungslose *und* optische Elektronenübergänge zu erwarten. Die Wärmeausbeutemessungen erfassen alle strahlungslosen Verluste. Sie können daher einen Anhaltspunkt für die Aufteilung des Prozesses II in strahlungslose und optische Elektronenübergänge liefern.

Während der Prozess I von chemischen Zusätzen im Kristall und von der Temperatur unabhängig ist, kann der Ablauf des II. Prozesses durch absichtlichen Einbau von Störstellen in den Kristall verschieden geleitet werden.

Beim *reinen* Kristall, bei dem, abgesehen vom ersten Moment der Verfärbung, alle gebildeten Paare rekombinieren, zeigt die Wärmeausbeute von 90%, dass die Elektronen-Defektelektronen-Rekombination etwa zur Hälfte strahlungslos erfolgt.

Das *hydrierte* KBr dient als Beispiel für einen Kristall, in dem zu Anfang der Bestrahlung die Rekombination durch eingebaute Fangstellen vollständig unterbunden wird. Durch den Einfang eines Paares werden hier etwa 4 eV photochemisch gespeichert. Das ist die Hälfte des Band–Band-Abstands und entspricht 10% der absorbierten Röntgenenergie. Aus der Anfangswärmeausbeute von 87% für den hydrierten Kristall kann man nun schliessen, dass die Übergänge vom Leitungs- bzw. Valenzband zur betreffenden Fangstelle grösstenteils strahlungslos erfolgen. Das Ergebnis verträgt sich gut mit der Beobachtung, dass hydrierte Kristalle bei Röntgenbestrahlung viel schwächer leuchten als reine Kristalle. Die Wärmeausbeute des hydrierten Kristalls nähert sich im Laufe der Verfärbung der des reinen, weil die Fangstellen allmählich verbraucht werden und dadurch ein allmählich immer grösserer Teil der Träger rekombiniert.

Leider lassen sich aus der Wärmeausbeute röntgenphotochemischer Prozesse nur ziemlich ungenaue Aussagen über die Anteile der optischen und strahlungslosen Elektronenübergänge gewinnen, da sich diese dem viel grösseren Energieverlust beim Prozess I überlagern und selbst nur wenige Prozente am ganzen Energieumsatz ausmachen.

Genauere Aussagen über den Anteil der optischen und strahlungslosen Elektronenübergänge lassen sich aus Wärmeausbeutemessungen bei *optischer* Elektronenanregung gewinnen, wo die Elektronenübergänge für sich allein auftreten. Als Beispiel dafür wurde die Wärmeausbeute beim Übergang angeregter F-Zentren in den Grundzustand gemessen. Die Messmethode entsprach der in § 2 für das röntgenphotochemische Problem beschriebenen. Die Wärmeausbeute ergab sich zu etwa 61% für −180°C und etwa 93% für −78°C. Gemessen wurde im Gleichgewichtszustand, in

* Es war für eine graphische Auswertung nach dem in § 4 beschriebenen Verfahren dabei jeweils eine 3,5 Minuten lange Bestrahlung nötig, anschliessend eine ebenso lange Dunkelpause.

dem keine photochemische Energiespeicherung erfolgt. Die restliche Energie muss also ausgestrahlt werden. Für den gleichen Prozess wurde von BECKER und PICK[6] die Fluoreszenzquantenausbeute gemessen. Ihre Ergebnisse bei gleicher F-Zentren-Konzentration ergänzen nach Umrechnung auf die energetische Ausbeute obige Werte bis auf wenige Prozent genau zu 100%, sodass sich die Messungen gegenseitig bestätigen. Dies kann gleichzeitig als Bestätigung des Wärmeausbeutemessverfahrens überhaupt gewertet werden.

Herrn Professor Dr. H. PICK danke ich herzlich für die Aufnahme in sein Institut, für die Anregung zu dieser Arbeit und für sein dauerndes reges Interesse.

Herrn Dr. F. LÜTY habe ich für viele wertvolle Diskussionen zu danken.

LITERATUR

1. LÜTY F. *Z. Phys.* **145,** 249 (1956).
2. POHL R. W. *Phys. Z.* **39,** 36 (1938).
3. SEITZ F. *Rev. Mod. Phys.* **18,** 384 (1946); *Ibid.* **26,** 7 (1954).
4. MARTIENSSEN W. und PICK H. *Z. Phys.* **135,** 309 (1953).
5. MIESSNER G. *Z. Phys.* **141,** 111 (1955).
6. BECKER K-H. und PICK H. *Nachr. Ges. Wiss. Göttingen* 167, (1956) und unveröffentlichte Messungen.

VOL.
4
1958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 58–67.

THE EFFECT OF ION PAIR AND ION TRIPLET FORMATION ON THE SOLUBILITY OF LITHIUM IN GERMANIUM—EFFECT OF GALLIUM AND ZINC

HOWARD REISS and C. S. FULLER

Bell Telephone Laboratories, Inc., Murray Hill, N.J.

(*Received* 15 *May* 1957)

Abstract—The present investigation was undertaken to test the prediction that ion pairing or association in solids should lead to increased solubility of a solute in equilibrium with an external phase. The effect of association between lithium ions in solution in germanium with substitutional gallium ions (singly charged) as well as with substitutional zinc (doubly charged) has been examined. Agreement with theory, including the pronounced effect of the doubly charged ions, has been demonstrated. The results also confirm the predicted minimum in the solubility–temperature curves caused by the mass action effect of the hole-electron equilibrium in germanium. It is suggested that solubility determinations may be useful in determining the number of ionization levels of impurities.

By taking account of ion-association effects a more accurate curve for the solubility of lithium in undoped germanium has been obtained.

1. INTRODUCTION

ION pairing in semiconductors is now a well-established fact[1,2] but no quantitative study of the relation between ion pairing and solubility has been performed. This omission is a result of the experimental difficulties which attend such an undertaking, and, to a lesser extent, of pessimism generated by limited calculations performed in Appendix A.[1] These calculations indicated that the *room temperature* solubility of lithium was only slightly dependent upon pairing, due to the overwhelming influence of the hole-electron equilibrium. However, subsequent estimates have shown that insensitivity is confined only to the neighborhood of room temperature. At higher (but not too high) temperatures easily observable effects should occur.

A truly definitive study of ion pairing and the solubility of lithium in *doped* germanium cannot be made without some knowledge of the solubility of lithium in *undoped* germanium. In Section IV[1] an attempt to measure this solubility (from an external phase formed by alloying lithium to germanium) was made. An indirect method was employed to overcome the difficulty that lithium diffused into germanium at an elevated temperature precipitates at a lower temperature, e.g. room temperature, to which the germanium specimen must be cooled for the purpose of performing electrical measurements. Precipitation occurs because the solubility, $N_D{}^0$, in undoped germanium decreases with decreasing temperature. An equilibrium solution formed at a higher temperature will therefore be supersaturated at a lower temperature.

The indirect method of determining $N_D{}^0$ consisted of diffusing lithium into gallium-doped germanium from which precipitation can be prevented by taking suitable advantage of the hole-electron equilibrium (cf.[1]). Electrical measurements at room temperature could then be used to determine the solubility in the doped germanium at the temperature of diffusion. This solubility could then be used in equation (4.1)[1] to calculate $N_D{}^0$ (denoted by $D_0{}^+$[1]). In this manner the dashed curve of Fig. 1, which is a reproduction of Fig. 5,[1] was obtained. For convenience hereafter we shall refer to this curve as the *original* solubility curve.

Now equation (4.1)[1] takes no account of ion pairing and therefore the reliability of the original solubility curve is open to question. Furthermore

the solubilities in gallium-doped specimens were obtained by diffusing lithium to saturation. At the lower temperatures, e.g., the 100°C point represented by an x in Fig. 1, the diffusivity of lithium is

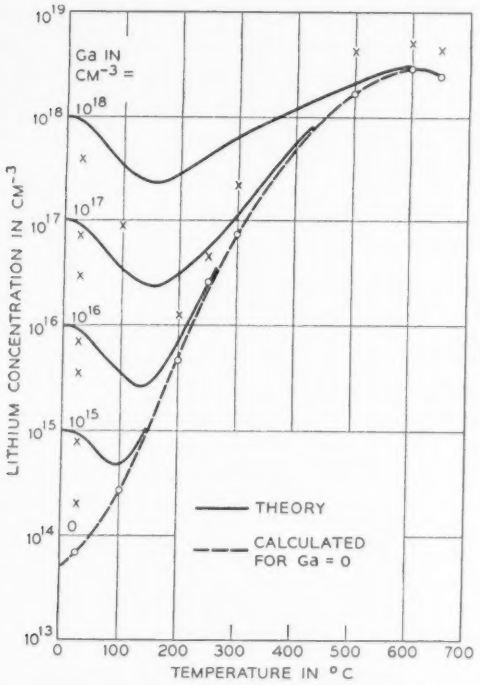


FIG. 1. Solubility curves for Li in Ge reproduced from REISS *et al.*[1] The solid curves, which give the solubility as a function of temperature for various Ga dopings, are calculated without taking account of ion pairing. The dashed curve gives the solubility of Li in pure Ge calculated from the determinations shown as crosses in the figure and also does not take account of ion pairing effects. The last experimental point at 650°C is somewhat low because of some precipitation during cooling.

so low that many months are required to achieve saturation. According to the available data[3] on the diffusivity of lithium in undoped germanium 6 months should have been required for saturating the 100°C specimen. This period of time was allowed in the experiment, but later work on ion pairing (cf. Fig. 24,[1]) demonstrated that the diffusivity in gallium-doped germanium was lowered by ion pairing at 100°C. Thus it is probable that the x-value in Fig. 1 is low, i.e. that the specimen was not really completely saturated, and as a result, that the value of $N_D{}^0$ computed from equation (4.1) is low at 100°C. (Further discussion of this situation will appear later.)

The present experiments have been devised to obviate the difficulties mentioned above and have made possible through the use of ion pair theory[4] the calculation of a revised $N_D{}^0$ curve. Furthermore, added confidence in this result, and therefore the demonstration of the influence of ion pairing on solubility, is given by an independent check of the $N_D{}^0$ curve by MORIN and REISS[1] while studying lithium precipitation in germanium.

The theory of ion pairing predicts that the solubility of lithium in zinc-doped germanium should differ from that in gallium-doped germanium, since zinc as a double acceptor[2,5] in germanium is doubly charged. Observation of this difference therefore provides another demonstration of the existence of ion pairing. (Studies of this kind will be described below.)

The curves for solubility in doped germanium depicted in Fig. 1 exhibit minima. Until now this interesting feature has remained a theoretical prediction. In the present article we shall demonstrate such minima experimentally.

2. EXPERIMENTAL

As mentioned above, the points marked by x's in Fig. 1 were obtained by attempting to diffuse lithium to saturation in germanium. In the present work such diffusions were also performed, but in addition a more expeditious method was employed. This consisted of saturating a specimen at an elevated temperature T_1 where the diffusivity is high, transferring it to a lower temperature T_2 where the solubility is less, and having it precipitate to equilibrium. In this manner the disadvantage of the low diffusivity at T_2 was overcome.

Rod shaped specimens $0{\cdot}5 \times 1{\cdot}5 \times 0{\cdot}12$ cm were cut from suitably doped zone-leveled ingots of germanium.* These were measured for resistivity by means of a two point probe device with precision of better than 1 per cent. Groups of specimens having as nearly as possible the same concentration of doping agent (gallium or zinc) were selected. The range of impurity concentrations within a group is indicated by resistivity values appearing in Tables 1 and 2 below.

Application of lithium was made from a lithium

* The authors are indebted to L. P. ADDA of the Allentown Laboratory for providing these ingots.

emulsion* applied by means of a camel's-hair brush to the large area sides of the germanium bars. Lithium was alloyed to the faces of the bars by passing a current of 6 A through them for 10 sec. The apparatus used is shown in Fig. 2. The maximum temperature of alloying was about 400°C.

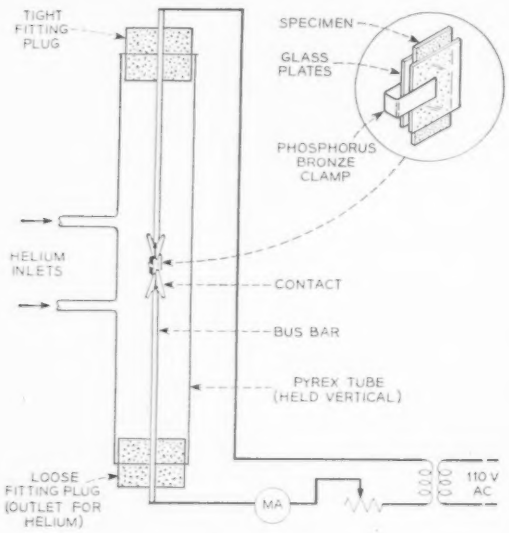


FIG. 2. Apparatus employed to alloy Li emulsion on specimen surfaces.

After alloying, the specimens were sealed in dry pyrex tubes under vacuum and equilibrated at various temperatures, (T_1), in the neighborhood of 300°C in a furnace controlled to ±1°C for 75–100 hr. Some of these specimens were further equilibrated at lower temperatures (T_2) for at least 48 hr. In this manner precipitations were caused to occur until the concentrations reached those corresponding to the solubilities at T_2, also controlled to ±1°C.

The fully equilibrated specimens were cooled rapidly to room temperature, the excess of lithium removed by grinding on abrasive paper, and their resistivities measured. The concentration of lithium (donor) added was then determined from the change in resistivity and the mobility† corrected for the new concentration of scattering centers. As a check lithium was removed by outward diffusion in air in a number of cases, whereupon the original resistivities were recovered to within the limit of measurement error in most cases.

3. THEORY

Before treating the further experimental details it will be advantageous to outline some of the theoretical concepts involved in handling the data.

(a) *Singly-charged acceptors.* In the case of *gallium*-doped germanium we shall be concerned with the following set of chemical equilibria.

$$\begin{array}{lccc} \mathrm{Li\,(alloy)} = & \mathrm{Li}^+ & + & \mathrm{e}^- \\ & + & & + \\ & \mathrm{Ga}^- & & \mathrm{e}^+ \\ & \| & & \| \\ & [\mathrm{Li}^+\mathrm{Ga}^-] & & \mathrm{e}^+\mathrm{e}^- \end{array} \qquad (3.1)$$

Here Li^+ and Ga^- represent ions dissolved in germanium, e^+ and e^-, are a valence band hole and conduction band electron respectively, and e^+e^- is a "recombined" hole-electron pair. As in reference (1) $[Li^+Ga^-]$ represents an ion pair formed between lithium and gallium. Li (alloy) stands for lithium dissolved in the external alloy.

For concentrations of the various species the following symbols will be used:

$$\left.\begin{array}{ll} D^+ & = \text{concentration of } Li^+ \\ A^- & = \text{concentration of } Ga^- \\ n & = \text{concentration of } e^- \\ p & = \text{concentration of } e^+ \\ P & = \text{concentration of } [Li^+\,Ga^-] \\ \alpha & = \text{thermodynamic activity of lithium in the alloy} \\ N_D & = \text{total concentration of lithium} \\ N_A & = \text{total concentration of gallium} \\ N_D^\circ & = \text{solubility of lithium in the absence of acceptor} \\ n_i & = \text{concentration of intrinsic electrons or holes.} \end{array}\right\} \qquad (3.2)$$

By definition

$$\begin{aligned} N_A &= A^- + P \\ N_D &= D^+ + P. \end{aligned} \qquad (3.3)$$

Besides these relations there exist mass action expressions corresponding to each of the equilibria in (3.1). Thus we have

$$D^+ n = \alpha K' = K(\text{for constant } \alpha), \qquad (3.4)$$

* This was supplied by the Foote Mineral Company and was first suggested by H. GUGGENHEIM of these Laboratories.

† Any effect of the precipitate on mobility should be to reduce it, thus leading to greater apparent solubility. However, no effects of any importance in this regard were noted.

VOL. 4 1958

for the donor equilibrium

$$\frac{P}{A^- D^+} = \Omega \tag{3.5}$$

for the ion pairing equilibrium, and

$$np = n_i^2 \tag{3.6}$$

for the hole-electron equilibrium.
To these is added the charge balance equation

$$D^+ + p = A^- + n. \tag{3.7}$$

Equations (3.3) through (3.7) can be manipulated to yield the following relations.

$$K = \frac{(N_D{}^0)^2}{2} + \sqrt{\left(\frac{(N_D{}^0)^4}{4} + n_i^2 (N_D{}^0)^2\right)} \tag{3.8}$$

$$\left(1 + \frac{n_i^2}{K}\right) D^+ = \frac{K}{D^+} + \frac{N_A}{1 + \Omega D^+} \tag{3.9}$$

$$N_D = D^+ \left(1 + \frac{\Omega N_A}{1 + \Omega D^+}\right). \tag{3.10}$$

According to (3.8), K can be obtained when n_i and $N_D{}^0$ are known. Ω can be calculated by the procedure outlined in Section IX[1] provided that a, the distance of closest approach between ions, is known. With K, Ω, n_i, and N_A known (3.9) can easily be solved numerically for D^+. The latter can be inserted in (3.10) for the specification of N_D. Thus N_D can be plotted as a function of temperature (for given N_A) if K, Ω, and n_i are available as functions of temperature.

Unfortunately, as stated above, $N_D{}^0$ is not immediately known, and one of the tasks of this research will be to determine it with the assistance of experimental data on N_D. Equation (3.10) can be solved for D^+,

$$D^+ = -\frac{1}{2}\left(N_A - N_D + \frac{1}{\Omega}\right) + \sqrt{\left[\frac{1}{4}\left(N_A - N_D + \frac{1}{\Omega}\right)^2 + \frac{N_D}{\Omega}\right]}. \tag{3.11}$$

With K unknown, this value of D^+ can be substituted in (3.9) and the latter equation easily solved for K. Finally, $N_D{}^0$ is computed by inverting (3.8).

$$N_D{}^\circ = \frac{K}{\sqrt{(K + n_i^2)}}. \tag{3.12}$$

(b) *Doubly-charged acceptors.* In the case of zinc-doped germanium, the following set of chemical equilibria are involved:

$$\begin{array}{rccc} \text{Li (alloy)} = & \text{Li}^+ & + & \text{e}^- \\ & + & & + \\ & \text{Zn}^= & & \text{e}^+ \\ & \| & & \| \\ & [\text{Li}^+\text{Zn}^=] & & \text{e}^+\text{e}^- \\ & + & & \\ & \text{Li}^+ & & \\ & \| & & \\ & [(\text{Li}^+)_2\text{Zn}^=]. & & \end{array} \tag{3.13}$$

Here $Zn^=$ is a doubly charged zinc ion dissolved in germanium, $[Li^+Zn^=]$ a singly (negative) charged ion pair, and $[(Li^+)_2Zn^=]$ a neutral ion triplet.

For concentrations of the new species the following symbols are introduced:

$$\begin{array}{lcl} A^= & = & \text{concentration of } Zn^= \\ P^- & = & \text{concentration of } [Li^+Zn^=] \\ T & = & \text{concentration of } [(Li^+)_2Zn^=]. \end{array} \tag{3.14}$$

The total concentration of zinc will be given by N_A. It will always be clear whether we are dealing with gallium or zinc.

By definition,

$$\begin{aligned} N_A &= A^= + P^- + T \\ N_D &= D^+ + P^- + 2T. \end{aligned} \tag{3.15}$$

Besides (3.15) there are the mass action relations

$$D^+ n = K \tag{3.4}$$

$$\frac{P^-}{A^= D^+} = \Omega_P \tag{3.16}$$

$$\frac{T}{P^- D^+} = \Omega_T \tag{3.17}$$

$$np = n_i^2. \tag{3.6}$$

The charge balance equation becomes now

$$D^+ + p = 2A^= + P^- + n. \tag{3.18}$$

Equations (3.16) and (3.17) are associated with the equilibria involving pairs and triplets. Ω_P and Ω_T are constants which can be evaluated by the methods described in reference 2.

The set (3.15) through (3.18) can be manipulated to yield

$$\left[D^+\left(1+\frac{n_i^2}{K}\right)-\frac{K}{D^+}\right](1+D^+\Omega_P+(D^+)^2\Omega_P\Omega_T). = N_A[2+D^+\Omega_P] \tag{3.19}$$

$$N_D = D^+\left(1+\frac{N_A\Omega_P+2D^+N_A\Omega_P\Omega_T}{1+D^+\Omega_P+(D^+)^2\Omega_P\Omega_T}\right). \tag{3.20}$$

With a knowledge of K, Ω_P, Ω_T, n_i, and N_A (3.19) can easily be solved numerically for D. This result can be inserted in (3.20) to get N_D.

4. GALLIUM RESULTS, DETERMINATION OF D^0 AND CONFIRMATION OF THE OCCURRENCE OF ION PAIRING

Table 1 lists the results of experiments performed with gallium-doped germanium. Part A exhibits data pertaining to specimens equilibrated by diffusion alone, while B refers to cases in which saturation occurred by diffusion at one temperature T_1, and equilibration was achieved by precipitation at another temperature T_2, where the solubility is lower. As is apparent from the original resistivities

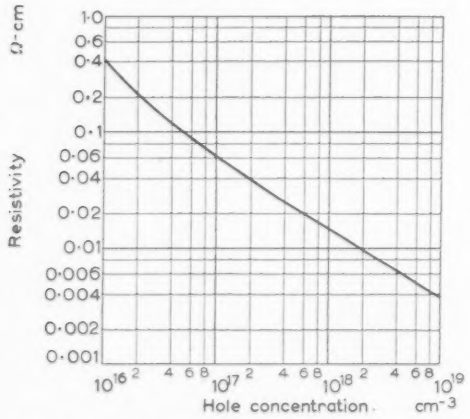


FIG. 3. Curve showing relation between resistivity of germanium and hole concentration.

ρ_0, the ingot was quite uniform in gallium. Lithium concentrations were calculated from the resistivity versus carrier concentrations presented in Fig. 3.*

Some adjustment in the values of N_D is required because of the minor fluctuations in gallium content of the various specimens. The last column in Table 1 lists lithium solubilities in atoms/cm^3 after a linear adjustment was made to make the solubility correspond to a gallium content of 4×10^{17} cm^{-3}. That is, if the specimen involved in the experiment contained only $3{\cdot}8\times10^{17}$ cm^{-3} gallium, then the measured solubility was multiplied by the factor 4/3·8. This linear adjustment is valid for small corrections.

Part B also contains data for two boron-doped specimens. These data are in line with the results on gallium-doped germanium, indicating that boron (as expected) has about the same effect as gallium.

Part C gives results for a series of specimens doped with gallium to lower levels than the specimens of Part A and B, i.e. to levels of about 2×10^{17} cm^{-3}.

In Fig. 4 results from Part A of Table 1 are plotted as circles while the data of Part B are plotted as x's. These, of course, are the results N_D^* rather than N_D. Curve A drawn through the data is an estimate of the best average result. The agreement between data from Parts A and B should be noted. Since equilibration was achieved by two different procedures the agreement is presumptive evidence for the attainment of equilibrium. The scatter, such as it is, is probably due to fluctuations in temperature or to incomplete precipitation.

In curve A, the unequivocal appearance of a minimum, as predicted by theory, should be noted. Using curve A for values of N_D we can now calculate $N_D{}^0$. To accomplish this, Ω, the ion-pairing equilibrium constant of equations (3.9), (3.10), and (3.11) must be available. In calculating

* Fig. 3 differs somewhat from other published results. It was obtained by compensating gallium-doped specimens, whose resistivities were known, with lithium, and then removing the lithium by diffusing it to the surface. This lithium was then determined by spectrometry. Since the specimens had been compensated the measured lithium determined the gallium content. The authors are indebted to E. JAYCOX and D. BABUSCI for these analyses.

VOL. 4 1958

Table 1. *Dependence of lithium solubility on temperature in germanium doped with gallium. Part A: Equilibrated by diffusion alone. Approximate gallium content* $= 4\times10^{17}$ cm^{-3}

T°C	ρ_0(Ωcm)	ρ(Ωcm)	$N_D\times10^{-17}$ (cm^{-3})	$N_D^*\times10^{-17}$ (cm^{-3})
240	0·0249	0·0319	1·2	1·2
240	0·0245	0·0329	1·4	1·4
245	0·0241	0·0297	1·2	1·2
245	0·0233	0·0299	1·4	1·3
245	0·0241	0·0297	1·1	1·1
245	0·0239	0·0299	1·3	1·3
250	0·0253	0·0328	1·2	1·3
298	0·0247	0·0374	1·8	1·9
300	0·0247	0·0347	1·6	1·6
300	0·0249	0·0349	1·5	1·5
300	0·0246	0·0405	2·1	2·1
300	0·0243	0·0394	2·1	2·1
300	0·0251	0·0361	1·6	1·7
300	0·0237	0·0374	2·1	2·0
318	0·0239	0·0470	2·6	2·5
360(†)	0·0240	0·0300(n)	5·0	4·9

* Corrected to 4×10^{17} cm^{-3} gallium.
(†) Converted to n-type.

Part B: Saturated at T_1 *and precipitated at* T_2. *Approximate gallium content* $= 4\times10^{17}$ cm^{-3}

T_1°C	T_2°C	ρ_0(Ωcm)	ρ(Ωcm)	$N_D\times10^{-17}$ (cm^{-3})	$N_D^*\times10^{-17}$ (cm^{-3})
300	98	0·0262	0·0606	2·5	2·8
300	150	0·0232	0·0324	1·7	1·6
325	150	0·0261	0·0380	1·5	1·7
325	150	0·0260	0·0415	1·8	2·0
325	200	0·0255	0·0384	1·8	1·9
325	200	0·0265	0·0346	1·2	1·4
360	100	0·0232	0·0659	3·3	3·1
360	250	0·0242	0·0320	1·4	1·4
360	100	0·0235	0·0687	3·3	3·1
493	318	0·0239	0·0590	3·1	2·9
389(†)	103	0·0255	0·0777	3·1	3·3
489(†)	248	0·0256	0·0324	1·1	1·2

* Corrected to 4×10^{17} cm^{-3} gallium.
† Boron doped.

Part C: Additional results by saturation and precipitation. Approximate gallium content $= 2\times10^{17}$ cm^{-3}

T_1°C	T_2°C	ρ_0(Ωcm)	ρ(Ωcm)	$N_D\times10^{-16}$ (cm^{-3})	$N_D^*\times10^{-16}$ (cm^{-3})
290	105	0·0374	0·0726	9·1	13·4
290	132	0·0394	0·0556	7·7	10·1
325	150	0·0392	0·0946	7·2	7·6
290	177	0·0396	0·0483	5·0	6·7
325	200	0·0409	0·0759	7·5	5·6
210	220	0·0399	0·0530	6·6	5·6
325	250	0·0391	0·0700	6·6	6·7
210	290	0·0505	0·0978	8·5	12·8

* Corrected to $2{\cdot}0\times10^{17}$ cm^{-3} gallium.

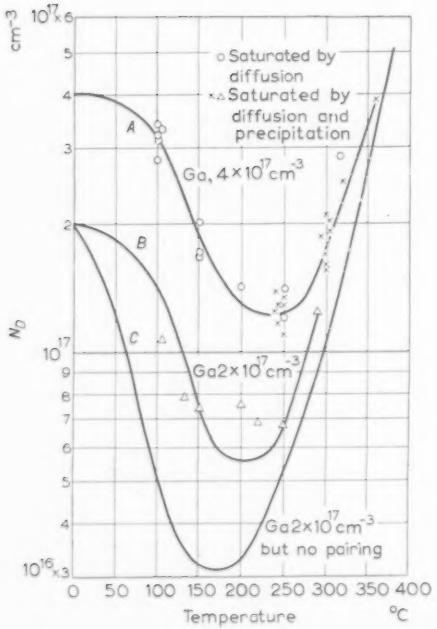


FIG. 4. Solubility of Li in Ge doped with Ga showing the reduction in solubility because of ion pairing. (A) Experimental points and theoretical curve for a Ga concentration of 4×10^{17} cm^{-3}. (B) Experimental points and theoretical curve for a Ga concentration of 2×10^{17} cm^{-3}. (C) Theoretical curve for Ga $= 2\times10^{17}$ cm^{-3} assuming no pairing present.

Ω by the methods described (which are based on an ion pair model in which the binding energy is coulombic) we assume a, the distance of closest approach between ions, to be $1{\cdot}71\times10^{-8}$ cm, a value obtained by HALL measurements on ion pairing between lithium and gallium in germanium.[1]

Taking N_D from curve A, Fig. 4, and $N_A = 4\times10^{17}$ cm^{-3}, equation (3.11) was employed to compute D^+. Insertion of the latter in (3.9) gave K. Then use of (3.12) together with known values[6] of n_i gave $N_D{}^0$. The results are shown in Fig. 5 as curve A, and are plotted as log $N_D{}^0$ versus reciprocal temperature.* The "original" solubility curve, also shown in Fig. 1, appears as curve B in Fig. 5. It does not differ much from A. This, in a sense, is fortunate since the work of MORIN and REISS[4] mentioned above has approximately validated the original solubility curve, obtained with neglect of ion pairing. Actually MORIN and REISS checked the original solubility curve at temperatures above 300°C ($1/T = 0{\cdot}00174$) and at these temperatures curves A and B are identical—as is expected since no ion pairing occurs at such high temperatures when only singly charged acceptors are present. They did obtain one check at room temperature ($1/T = 0{\cdot}00333$). This result appears as the filled circle in Fig. 5. It represents at least an upper limit to $N_D{}^0$ at room temperature, for the precipitation process leading to it required many weeks and may not have been complete when the final measurement was made.

* It should be emphasized that although $N_D{}^0$ is obtained through a consideration of results involving ion pairing, $N_D{}^0$ itself is independent of pairing effects.

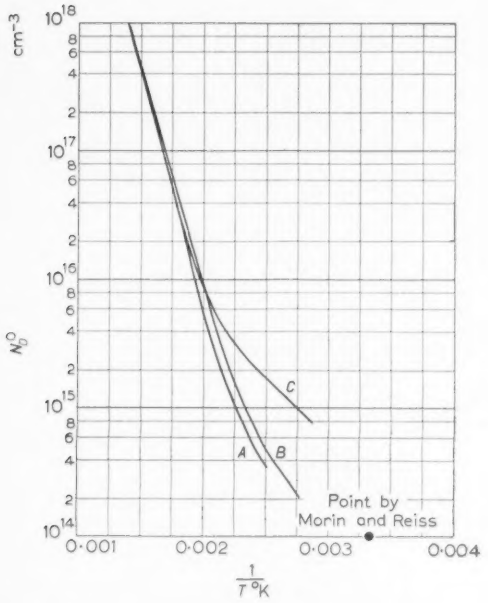


FIG. 5. Solubility Curves for Li in pure Ge. (A) Revised to take account of ion pairing. (B) Previous curve from[1] calculated without consideration of ion pairing. (C) Curve calculated from data of Fig. 4 curve A but with neglect of ion pairing.

It is natural to inquire as to why the inclusion or neglect of ion pairing makes so little difference between curves A and B. The answer is that the inclusion of pairing does make a difference but that the original curve B was calculated from N_D values some of which were in error. This is particularly true of the 100°C point mentioned in Section 1. To emphasize this, $N_D{}^0$ was computed using the data of curve A, Fig. 4, but with neglect of ion pairing. Curve C of Fig. 5 is obtained. It departs appreciably from curve A at low temperatures where ion pairing is reasonably complete. Since the room temperature point of MORIN and REISS lies below curve C, and that point is an *upper* limit, curve C is in obvious disagreement with it. On the other hand, curve A is in good agreement, and this constitutes an independent check (by solubility means) that ion pairing takes place.

It is now clear that the 100°C specimen of Fig. 1 was not saturated. This follows from the fact that although the gallium content was about 2×10^{17} cm^{-3}, the solubility of lithium appeared to be $0{\cdot}9\times10^{17}$ cm^{-3} (see Table 2).[1] From curve B of Fig. 4 in this article which relates to a gallium content of 2×10^{17} cm^{-3} we see that the lithium solubility at 100°C should have been about $1{\cdot}5\times10^{17}$ cm^{-3} rather than $0{\cdot}9\times10^{17}$ cm^{-3}. Curve A of Fig. 5 must now replace curve B as the most reliable lithium solubility curve for undoped germanium.

Using curve A of Fig. 5 together with (3.8), (3.9), and (3.10) curve B of Fig. 4 was computed, referring to a gallium content, N_A, of 2×10^{17} cm^{-3}. The triangles represent experimental values of N_D taken from the last column of Part C, Table 1. Aside from those at 200°C and 220°C the other points fall fairly well on the theoretical curve. The points at 200°C and 220°C, lying high as they do, probably represent incomplete precipitations. In general, however, considering the experimental difficulties, the agreement is satisfactory.

Curve C of Fig. 4 is for N_D when $N_A = 2\times10^{17}$ cm^{-3}, and when $\Omega = 0$, i.e., when no ion pairing is involved. We see that ion pairing makes a large difference, but not so much in the neighborhood of room temperature, a fact already noted.[1]

5. RESULTS WITH ZINC-DOPED GERMANIUM REFERENCES

The resistivity data for zinc-doped germanium require special treatment before they can be translated into concentrations. Zinc is a double acceptor[5,2] and at room temperature (where resistivity is measured) its deep level at 0·09 eV is not completely ionized at the zinc concentrations employed in our work. Furthermore, the doubly charged zinc ions affect the carrier (hole) mobility differently than do equal concentrations of gallium ions. The standard concentration–resistivity curves (Fig. 3) cannot therefore be used without certain adjustments.

These difficulties were overcome by the introduction of some fairly involved but nevertheless valid procedures. Our results with zinc, do confirm the theory in a semiquantitative manner. We shall omit a lengthy description of the adjustment procedures, and merely point out that they cannot be responsible for the failure of the data to agree more closely with theory.

After the zinc and lithium contents had been

Table 2. Temperature dependence of lithium solubility in zinc-doped germanium saturated at T_1 and precipitated at T_2. Approximate zinc content $= 1\times10^{17}$ cm^{-3}

T_1°C	T_2°C	ρ_0(Ωcm)	ρ(Ωcm)	$N_D\times10^{-17}$ (cm^{-3})	$N_D^*\times10^{-17}$ (cm^{-3})
290	105	0·0434	0·0726	1·35	1·61
309	105	0·0500	0·994	1·24	1·92
290	132	0·0394	0·0556	1·89	1·80
309	132	0·0503	0·277	1·13	1·76
325	150	0·0392	0·0946	1·48	1·46
309	175	0·0517	0·127	0·85	1·39
290	177	0·0396	0·0483	1·68	1·60
325	200	0·0409	0·0759	1·10	1·19
309	218	0·0500	0·094	0·73	1·13
290	220	0·0399	0·0530	1·58	1·52
325	250	0·0391	0·0700	1·13	1·15
290	290	0·0978	0·850	1·07	1·66

* Corrected to 1×10^{17} cm^{-3} zinc.

estimated by these means they were adjusted (like the gallium results) linearly to a zinc content of 1×10^{17} cm^{-3}. This corresponds to a potential carrier (hole) content of 2×10^{17} cm^{-3} and standardizes the data for comparison with curve B of Fig. 4. Table 2 collects the results. Here T_1 and T_2 are the temperatures of diffusion and precipitation respectively, just as in Table 1.

Calculated values of N_D, for $N_A = 1\times10^{17}$ cm^{-3}, can be arrived at by applying equations (3.19) and (3.20). Both Ω_P and Ω_T were computed by the methods described in MORIN and REISS.[2] In this connection model 1,[2] the so-called "polarized" model was chosen for the calulation of Ω_T. The distance of closest approach was taken as $2{\cdot}76\times10^{-8}$ cm.[2] For Ω_P $1{\cdot}76\times10^{-8}$ cm was taken to be the distance of closest approach. This is obtained by adding the tetrahedral covalent radius of zinc[7] to the ionic radius of lithium, and reducing the result by about 10 per cent. This method of estimate is justified by empirical evidence gathered in connection with gallium, aluminum, and indium.[1]

With these values of Ω_P and Ω_T curve A of Fig. 6 is calculated. On the same graph, for comparison, we show curve B of Fig. 4. The x's in the Fig. are the data of Table 2. The agreement with curve A is only qualitative but the bulk of the points lie above the curve. This raises the suspicion that precipitation in these samples is incomplete—a strong possibility because the effect of ion pairing in depressing the diffusivity of lithium is pronounced even at high temperatures when doubly charged ions such as zinc are present (see Fig. 24).[1] In support of this idea we note that above 200°C, where the reduction in diffusivity is less

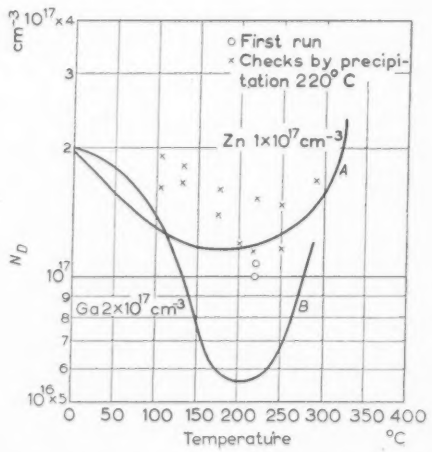


FIG. 6. Comparison of ion pairing effects of singly (Ga^-) and doubly ($Zn^=$) charged impurity ions. (A) Theory and experiment for Ge containing $1{\cdot}0\times10^{17}$ Zn per cm^3. Points marked by ○ are those shown by crosses at 100°C which were re-equilibrated at 220°C. (B) Theoretical curve for $2{\cdot}0\times10^{17}$ Ga cm^{-3} taken from Fig. 4.

pronounced, the x's lie closer on the average to the theoretical curve.

To check this possibility several specimens which, at lower temperatures, gave x's lying far above the curve were heated in vacuum at 220°C for 16 hr. In these specimens a small diffusion of lithium out of the germanium occurred during heating so that any results must be considered as lower limits to the true solubility at 220°C. These results appear as open circles in Fig. 6 and are in reasonable agreement with the theoretical curve.

Acknowledgements—The authors wish to thank A. J. PIETRUSZKIEWICZ JR. and F. J. DOLEIDEN for assistance with the experimental work.

OL.
4
958

REFERENCES

1. REISS H., FULLER C. S., and MORIN F. J. *Bell Syst. Tech. J.* **35,** 535 (1956).
2. MORIN F. J. and REISS H. *Phys. Rev.* **105,** 3 and 4 (1957).
3. FULLER C. S. and SEVERIENS J. C. *Phys. Rev.* **95,** 21 (1954).
4. MORIN F. J. and REISS H. *J. Phys. Chem. Solids* **3,** 196 (1957).
5. TYLER W. W. and WOODBURY H. H. *Bull. Amer. Phys. Soc.* **30,** 32 (1955).
6. MORIN F. J. and MAITA J. P. *Phys. Rev.* **94,** 1525 (1954).
7. PAULING L. *The Nature of the Chemical Bond* p. 179. Cornell (1942).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 68–70.

ION PAIRING IN SILICON

J. P. MAITA

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

(*Received* 25 *June* 1957)

Abstract—The occurrence of ion pairing in silicon is demonstrated. The pairing process is used to determine the diffusivity of lithium in silicon at low temperatures and the distance of closest approach between the ions forming the pair.

VOL. 4 1958

INTRODUCTION

THE formation of ion pairs in germanium has been discussed in detail in a recent paper by REISS, FULLER and MORIN.[1] Their work demonstrated the existence of ion pairs formed between Li^+, which is mobile, and Group III acceptors, A^-, which are fixed at ordinary temperatures in germanium. The investigations reported here deal with similar ion pairing experiments in silicon. Studies of relaxation time (τ), the distance of closest approach (a) between the ions forming a pair, and the diffusion constant (D) are discussed.

The diffusivity of lithium in silicon has been determined by FULLER[2] by measuring the motion of a p–n junction as lithium is diffused into a p-type sample. More recently, FULLER and SEVERIENS[3] have found an alternative method capable of greater experimental accuracy and based on fewer assumptions. This method consists of measuring the motion of p–n junctions as lithium drifts in an applied electric field. The results were in good agreement with those determined by the first method.

In this paper ion pairing has been used to determine diffusivity from measurements of relaxation times. These results extend the values of D from 10^{-8} cm^2/sec, the limit of the previous experiments, to 10^{-16} cm^2/sec.

EXPERIMENTAL

The specimens used in this investigation were cut from adjacent sections of a single crystal of floating-zone silicon doped with aluminum. This material was prepared by the floating-zone method to avoid the effects of heat treatment which occur in pulled crystals. Lithium was diffused into one sample from a tin–lithium bath[1] held at 445°C for 16 hr. Experiments were first performed in order to establish the correct temperature and lithium concentration required to avoid conversion from p- to n-type and yet yield a nearly compensated specimen. Another sample was used as a control. Both specimens were cut into bridges[4] and sand-blasted. Contacts were made by bonding through gold-plated areas with gallium-doped gold wire.

(*a*) *Carrier mobility*

Hall mobilities of both samples were determined from measurements of conductivity and Hall effect. The control sample contained $7{\cdot}7\times10^{16}$ acceptors per cm^3. After the diffusion of lithium, the sample to be used for ion pairing studies had a net acceptor concentration of $1{\cdot}46\times10^{16}$ per cm^3.

Fig. 1 shows plots of the logarithm of Hall mobility for both samples versus the logarithm of temperature in °K. Curve A is for the control and B for the lithium-doped specimen. As was anticipated, curve B lies above A, indicating the removal of charged impurity scatterers due to the formation of ion pairs. In obtaining these data, sample B was held at the experimental temperature until the pairing equilibrium had been achieved.

It has been shown[1] that at the temperature at which these curves cross (the "cross-over temperature") P/N_D is 0·5, where P is the concentration of ion pairs and N_D is the donor concentration. The best straight lines through the measured mobilities of samples A and B give a cross-over temperature of 470°K.

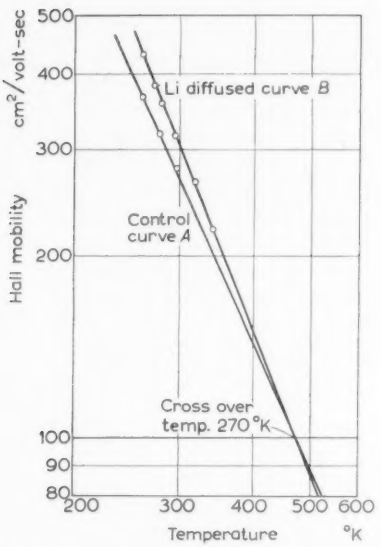


FIG. 1. Illustration of cross-over phenomenon for ion pairing in silicon. Curve A is the control with a concentration of $7{\cdot}7\times10^{16}$ cm^{-3} aluminum. Curve B represents a sample containing the same amount of aluminum to which lithium has been added.

REISS, FULLER and MORIN[1] have shown how to calculate the distance of closest approach (a) between two ions of opposite sign. They give an equation ((9.6) of reference 1):

$$\mathrm{Log}_{10}Q(\alpha) = \log_{10}\left[\left\{\frac{1}{4\pi}\left(\frac{\kappa kT}{q^2}\right)\right\}^3 \cdot \frac{\theta}{(N_A-\theta N_D)(1-\theta)}\right] \quad (1)$$

Here $Q(\alpha)$ is an intermediate function calculated in reference 1, while α is related simply to the distance of closest approach (a) by

$$a = \frac{q^2}{\kappa kT\alpha} \quad (2)$$

θ is P/N_D (the fraction of donors paired), T the cross-over temperature κ the dielectric constant (12 for silicon), k Boltzmann's constant, and N_A and N_D the concentration of acceptors and donors, respectively.

When $\theta = P/N_D$ has been determined, its value can be used in equation (1) to fix $Q(\alpha)$. The latter in turn can be compared with Table III of reference 1 (a tabulation of $Q(\alpha)$ versus α) to yield α by interpolation and therefore (a) by equation (2).

The experimental data give $a = 2{\cdot}0\times10^{-8}$ cm for lithium–aluminum in silicon. This distance was found to be $1{\cdot}68\times10^{-8}$ cm in germanium. The predicted value of $a = 1{\cdot}86\times10^{-8}$ cm, obtained by adding the covalent radius of Al to the ionic radius of Li^+, is very close to these experimental results. Because of the long extrapolation of the mobility curves, the result for silicon is uncertain by about $0{\cdot}3\times10^{-8}$ cm. The distance of closest approach does not enter into the relaxation calculations, which will be discussed below.

(*b*) *Relaxation*

Relaxation times were measured for the lithium-doped bridge (sample B). This bridge was heated at 373°K for approximately 30 sec and quenched quickly in liquid nitrogen (77·4°K). The heating dissociates the ion pairs and the low temperature inhibits pairing, so that ion pair formation can then be studied at some higher temperature. Ion pairing in sample B was studied at a temperature of 273·2°K, which was obtained by the use of a constant-temperature bath. This procedure was repeated at 300·4°K. At both temperatures ion pair formation takes place at a reasonable rate. The change in sample conductivity (as pairing took place) was measured as a function of time. Figs. 2 and 3 show plots of $\log(\sigma_\infty-\sigma)$ versus time. At both temperatures the curves are straight lines, demonstrating that ion pairing is a first-order reaction. The relaxation time (τ) can be computed from the equation

$$\tau = -\frac{\Delta t}{2{\cdot}3\Delta\log(\sigma_\infty-\sigma)} \quad (3)$$

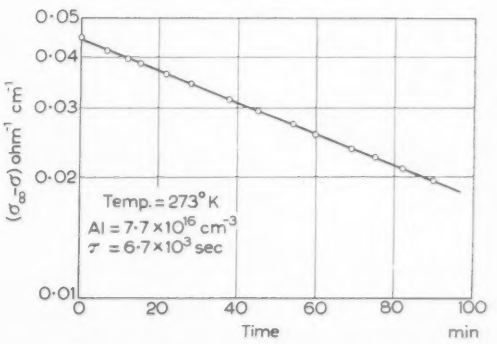


FIG. 2. Plot of log $(\sigma_\infty-\sigma)$ as a function of time, showing first-order kinetics of pairing at 273°K.

where σ_∞ is the equilibrium conductivity and σ is conductivity at time t.

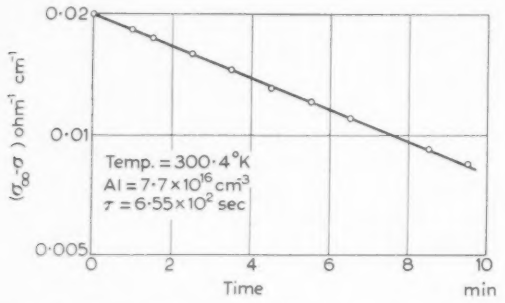


FIG. 3. Plot of log $(\sigma_\infty - \sigma)$ as a function of time, showing first-order kinetics of pairing at 300·4°K.

(c) *Diffusivity*

The data in Figs. 2 and 3 can be used to calculate the diffusivity of lithium from the equation[1]

$$D_0 = \frac{\kappa k T}{\tau 4\pi q^2 N}. \qquad (4)$$

The diffusivity was calculated at these temperatures and the results are shown in Fig. 4. This plot also shows the measurements obtained by FULLER and SEVERIENS.[3] These authors were able to obtain diffusivities of 10^{-5} to 10^{-8} cm^2/sec over a temperature range from 1150° to 633°K. The additional points obtained by relaxation now extend the curve from 10^{-8} to 10^{-16} cm^2/sec and also extend the temperature down to 273°K.

These results and those previously reported by FULLER and SEVERIENS both fall on the same straight line. The value obtained for the activation energy of diffusion over the temperature range from 1150° to 273°K is 16,700 cal/mole, and the diffusivity is given by the equation

$$D = 23{\cdot}0 \times 10^{-4} \exp(-16700/RT). \qquad (5)$$

Although these data were obtained at much lower temperatures, this result demonstrates again[1] the validity of the ion pairing concept and the ideas derived from it.

Acknowledgements—The author wishes to thank H. J. GUGGENHEIM for the difficult task of preparing the lithium specimen and F. J. MORIN for suggesting the problem and for helpful discussions.

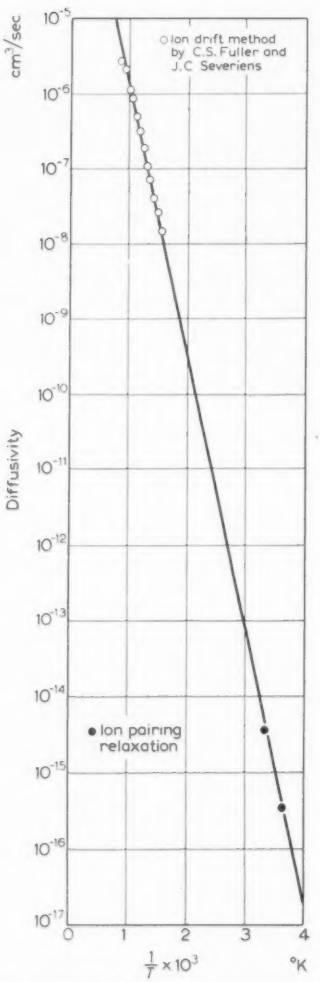


FIG. 4. Plot of diffusivity versus reciprocal of absolute temperature for lithium in silicon.

Thanks are due to H. REISS for many helpful comments during the course of the work.

REFERENCES

1. REISS H., FULLER C. S. and MORIN F. J. *Bell Syst. Tech. J.* **35**, 535 (1956).
2. FULLER C. S., THEUERER H. C. and VAN ROOSBROECK W. W. *Phys. Rev.* **85**, 678 (1952). FULLER C. S. *Phys. Rev.* **86**, 136 (1952).
3. FULLER C. S. and SEVERIENS J. C. *Phys. Rev.* **95**, 21 (1954).
4. MORIN F. J., GEBALLE T. H. and HERRING C. *Phys. Rev.* **105**, 525 (1957).

VOL. 4 1958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 71–77.

ANOMALIES DE RÉSISTIVITÉ DANS CERTAINS MÉTAUX MAGNÉTIQUES

P. G. DE GENNES* et J. FRIEDEL†

* Centre d'Etudes Nucléaires de Saclay; † Ecole Nationale Supérieure des Mines, Paris

(*Received* 27 *May* 1957; *revised* 17 *June* 1957)

Résume—La résistivité anormale de certaines terres rares et des alliages tels que AuMn, Au_3Mn, est interpretée en admettant qu'il existe un couplage d'échange entre électrons de conduction et porteurs de magnétisme. La valeur absolue de la section efficace correspondante est traitée comme une quantité phénomènologique. Aux températures élevées, les spins atomiques sont orientés au hasard, et les électrons de conduction ont un libre parcours moyen fini. Aux basses températures, les moments atomiques sont tous alignés et ne diffusent plus les électrons. Les effets d'ordre à courte distance sont analysés dans les limites de l'approximation de BORN. Ils sont petits dans la plupart des cas physiques.

Abstract—The anomalous resistivity of some rare earth metals, and of alloys like AuMn, Au_3Mn, is studied by assuming a coupling between conduction electrons and atomic spins. The magnitude of the corresponding cross section is treated as a phenomenological quantity. At high temperatures, the atomic spins are at random, and the conduction electrons have a finite mean free path. At low temperatures, the atomic spins are all alined and no scattering can occur. Short-range order effects in the spin lattice are analysed in the Born approximation, and shown to be small in most physical situations.

1. GÉNÉRALITÉS

ON connait, en fonction de la température, la résistivité du Gadolinium, du Dysprosium et de l'Erbium.[1] Ces substances sont soit ferromagnétiques soit antiferromagnétiques. Dans tous les cas, leur résistivité décroit brutalement en dessous du point de transition T_c.

Pourtant, dans ces substances, les porteurs du courant et les porteurs du magnétisme sont assez nettement *distincts*. Les premiers sont convenablement décrits par l'approximation des bandes. Les seconds, qui sont dans les niveaux *f* ou *d* fortement liés, sont vraisemblablement justiciables du modèle de ferromagnétisme de HEISENBERG. Ceci est bien verifié expérimentalement, notamment pour le Gadolinium : le nombre de porteurs par atome est *entier*[7] et l'intégrale d'échange $2J$ a une valeur bien définie, indépendante de la méthode de mesure. (Loi d'aimantation à basse température, chaleur spécifique.)

Par contre, des métaux comme le nickel, où les électrons doivent être traités par l'approximation des bandes,[4] sont exclus du traitement que nous allons développer. Ils ont du reste été étudiés du point de vue de la conductivité par MOTT.[5] Le cas des alliages or-manganèse est moins tranché; il semble cependant que l'on puisse attribuer aux atomes de manganèse un spin assez bien défini (compris entre 2–5/2).

Nous étudierons donc la conductivité avec le modèle suivant :

(a) des ions magnétiques de spin S répartis aux nœuds R d'un réseau rigide. On désignera par d la distance entre ions voisins, par v le nombre de voisins d'un ion donné, par c le nombre d'ions magnétiques par cm^3, par $2J\ \boldsymbol{S}_R \cdot \boldsymbol{S}_{R'}$ l'énergie d'échange de deux voisins.

(b) des électrons de conduction (ξ électrons par ion magnétique) traités comme des électrons libres, avec une masse m et un niveau de Fermi

$$\frac{\hbar^2 k_0^2}{2m}.$$

Enfin, on décrira le couplage d'échange entre ions et électrons libres par la forme simplifiée :

$$\mathscr{H}_I = \sum_{R,p} G\delta(\boldsymbol{R}-\boldsymbol{R}_p)\boldsymbol{S}_R \cdot \boldsymbol{S}_p \qquad (1.1)$$

S_p = spin du p^e électron.

Une telle interaction ne prétend pas représenter fidèlement les détails de la diffusion par un ion magnétique. Mais les résultats que nous obtiendrons ci-dessous sont, pour une bonne part, indépendants de tels détails.

En présence d'une interaction telle que (1.1) les spins ioniques, s'ils ne sont pas tous alignés, limitent le libre parcours moyen des électrons.* D'après les courbes expérimentales de résistivité, cet effet est au moins du même ordre de grandeur que celui des collisions avec phonons, pour des températures de l'ordre de T_c. Il est alors possible de soustraire, au moins approximativement, la partie de la résistivité dûe aux vibrations du réseau. Dans la suite de ce travail, nous nous intéressons uniquement à l'effet du désordre de spin.

2. DIFFUSIONS QUASI-ÉLASTIQUES

Nous nous proposons de montrer tout d'abord que les transferts d'énergie entre électrons et réseau des spins peuvent être négligés, sauf aux basses températures, dans le domaine d'existence des ondes de spin.

Plaçons un électron *unique* dans la bande de conduction, se déplacant avec la vitesse $\hbar k_0/m$. Lors d'un choc avec les spins ioniques, il cède l'énergie $\hbar\omega$. La moyenne quadratique peut être déduite de calculs analogues, qui ont été faits ailleurs[6,7] à propos de la diffusion magnétique des neutrons :

$$\langle\omega^2\rangle_T = \left\langle\left(\frac{d\boldsymbol{S}}{dt}\right)^2\right\rangle_T [S(S+1)-(\langle\boldsymbol{S}\rangle_T)^2]^{-1} \qquad (2.1)$$

le symbole $\langle\ \rangle_T$ représente une moyenne sur les états du réseau de spins à la température T. $d\boldsymbol{S}/dt$ est la vitesse de rotation d'un spin ionique *dûe au couplage ion-ion*. La formule n'est correcte que pour les grands transferts d'impulsion ($\hbar\kappa \gtrsim \pi\hbar/d$) mais elle donne dans tous les cas une limite supérieure approchée de $\langle\omega^2\rangle_T$. En négligeant les corrélations entre spins ioniques, autres que celles dûes à l'existence d'une aimantation spontanée (ce qui est analogue à une approximation de champ moléculaire) on déduit de l'hamiltonien d'échange ion ion la valeur approchée de $\langle(d\boldsymbol{S}/dt)^2\rangle$

$$\left\langle\left(\frac{d\boldsymbol{S}}{dt}\right)^2\right\rangle_T = \left(\frac{2J}{\hbar}\right)^2 V\left[\frac{S(S+1)-\langle S_z^2\rangle)^2}{2} - \frac{\langle S_z\rangle^2}{2} + 2[S(S+1)-\langle S_z^2\rangle]\langle S_z^2\rangle + (V-1)\times \times[S(S+1)-\langle S_z^2\rangle]\langle S_z\rangle^2\right] \qquad (2.2)$$

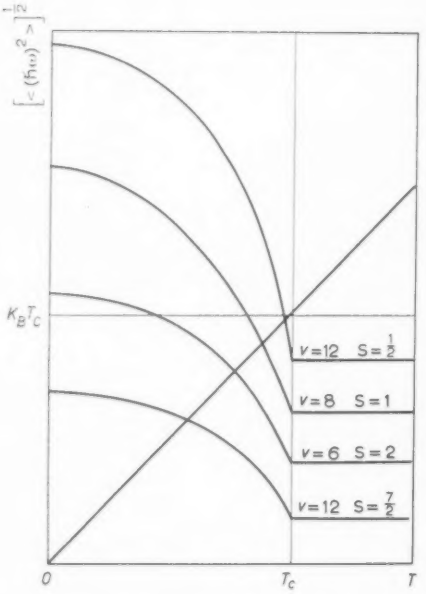


FIG. 1. Moyenne quadratique $\hbar\omega_{mq}$ de l'énergie échangée lors de la collision d'un électron isolé avec le réseau de spins. La ligne droite correspond à l'énergie thermique k_BT.

(Par convention, Oz désigne la direction de l'aimantation spontanée). Pour déterminer $\langle S_z\rangle_T$ à chaque température, on utilisera l'approximation de champ moléculaire. Pour $\langle S_z^2\rangle$ on prendra le modèle simple d'un spin fluctuant parmi des voisins à aimantation rigide. Les résultats sont résumés sur la figure 1. A part la région des basses températures, qui sera étudiée séparément, la condition :

* *Note added in proof :* ceci a été remarqué en premier lieu par T. KASUYA, Progress of Theoretical Physics **16**, 58 (1956).

VOL. 4 195

$$\langle(\hbar\omega)^2\rangle_T < k_B T$$

est approximativement satisfaite (d'autant mieux que le spin S est plus grand). En première approximation, les collisions sont *élastiques.*

3. SECTION EFFICACE DANS L'APPROXIMATION ÉLASTIQUE

Il nous suffit alors de connaître la section efficace integrée sur les énergies finales possibles de l'électron. Elle vaut dans l'approximation de BORN :

$$\frac{d\sigma}{d\Omega} = \left(\frac{m}{2\pi\hbar^2}\right)^2 \sum_{\alpha,\beta} w_\alpha \left| \langle\alpha| \sum_R \exp(i\boldsymbol{\varkappa}\,.\,\boldsymbol{R}) G^0 \boldsymbol{S}_p\,.\,\boldsymbol{S}_R - \boldsymbol{S}_p\,.\,\langle\boldsymbol{S}_R\rangle |\beta\rangle \right|^2 \tag{3.1}$$

$|\alpha\rangle$ et $|\beta\rangle$ désignent les états de spin du système total avant et après collision. w_α est la probabilité pour trouver le système dans l'état $|\alpha\rangle$ à l'équilibre thermique. $\hbar\boldsymbol{\varkappa}$ est la variation d'impulsion de l'électron.

La sommation sur les états finaux de spin $|\beta\rangle$ peut être effectuée et l'on obtient :

$$\frac{d\sigma}{d\Omega} = \left(\frac{m}{2\pi\hbar^2}\right)^2 \sum_\alpha w_\alpha \left\{ \sum_{R,R'} G^2 \exp[i\boldsymbol{\varkappa}\,.\,(\boldsymbol{R}-\boldsymbol{R}')] \langle\alpha| (\boldsymbol{S}_p\,.\,\boldsymbol{S}_R - \boldsymbol{S}_p\,.\,\langle\boldsymbol{S}_R\rangle)(\boldsymbol{S}_p\,.\,\boldsymbol{S}_{R'} - \boldsymbol{S}_p\,.\,\langle\boldsymbol{S}_{R'}\rangle) |\alpha\rangle \right\} \tag{3.2}$$

On a d'ailleurs

$$\begin{aligned} &\sum_\alpha w_\alpha \langle\alpha| \ldots |\alpha\rangle \equiv \langle \ldots \rangle_T \\ &= \langle (\boldsymbol{S}_p\,.\,\boldsymbol{S}_R - \boldsymbol{S}_p \langle \boldsymbol{S}_R\rangle)(\boldsymbol{S}_p\,.\,\boldsymbol{S}_{R'} - \boldsymbol{n}_p\,.\,\langle\boldsymbol{S}_{R'}\rangle)\rangle \\ &= \langle (\boldsymbol{S}_p\,.\,\boldsymbol{S}_R)(\boldsymbol{S}_p\,.\,\boldsymbol{S}_{R'})\rangle - \\ &\qquad - \langle (\boldsymbol{S}_p\,.\,\langle\boldsymbol{S}_R\rangle)(\boldsymbol{S}_p\,.\,\langle\boldsymbol{S}_{R'}\rangle)\rangle \\ &= \tfrac{1}{2} S_p(S_p+1) \times \\ &\qquad [\langle \boldsymbol{S}_R\,.\,\boldsymbol{S}_{R'}\rangle - \langle\boldsymbol{S}_R\rangle\,.\,\langle\boldsymbol{S}_{R'}\rangle]. \end{aligned} \tag{3.3}$$

Introduisons la section efficace incohérente par ion magnétique.

$$\sigma_0 = \frac{1}{4\pi}\left(\frac{mG}{\hbar^2}\right)^2 S(S+1). \tag{3.4}$$

La section efficace différentielle vaut, par ion magnétique :

$$\frac{d\sigma}{d\Omega} = \frac{\sigma_0}{4\pi} \sum_R (\langle \boldsymbol{S}_0\,.\,\boldsymbol{S}_R\rangle - \langle\boldsymbol{S}\rangle^2) \exp(i\boldsymbol{\varkappa}\,.\,\boldsymbol{R}) [S(S+1)]^{-1}. \tag{3.5}$$

La résistivité ρ vaut

$$\rho = \frac{m}{e^2 \tau c \xi} \tag{3.6}$$

τ est le temps de relaxation, défini ici par

$$\frac{1}{\tau} = c\frac{\hbar k_0}{m} \times 2\pi \int_0^\pi \sin\theta\, d\theta (1-\cos\theta) \frac{d\sigma}{d\Omega}. \tag{3.7}$$

Dans la limite des hautes températures (diffusion totalement incohérente), (3.7) prend la valeur :

$$\frac{1}{\tau_0} = c\frac{\hbar k_0}{m}\sigma_0. \tag{3.8}$$

4. APPROXIMATION SANS CORRÉLATIONS

D'après (3.5) le calcul de ρ suppose la connaissance des corrélations de spin

$$\gamma_R = [\langle \boldsymbol{S}_0\,.\,\boldsymbol{S}_R\rangle - \langle\boldsymbol{S}\rangle^2][S(S+1)]^{-1} \tag{4.1}$$

Toutefois, comme nous le montrerons plus loin, la longueur d'onde des électrons au niveau de Fermi est en général suffisamment courte pour que l'on puisse négliger les effets d'interférence entre ions magnétiques *differents.** On attend donc une

* Ceci a été remarqué par KORRINGA[16] pour un problème analogue.

diffusion *isotrope* avec la section efficace :

$$\frac{d\sigma}{d\Omega} = \sigma_0 \frac{S(S+1) - \langle \boldsymbol{S} \rangle^2}{S(S+1)} \qquad (4.2)$$

d'où l'on tire :

$$\frac{\tau_0}{\tau} = 1 - \frac{\langle \boldsymbol{S} \rangle^2}{S(S+1)}. \qquad (4.3)$$

Cette formule fait clairement apparaître le lien entre le temps de relaxation et l'ordre à longue distance dans le réseau des spins ioniques. Elle est comparable à la formule approchée de NORDHEIM pour les solutions solides désordonnées.

Les prévisions numériques, déduites de (4.3) au moyen d'une approximation de champ moléculaire, pour le calcul de $\langle \boldsymbol{S} \rangle$,* sont illustrées sur la figure 2. Au dessus de T_c, la résistivité dûe au désordre de spin est *constante* : ceci suggère une extrapolation simple pour retrancher la contribution des phonons (qui est ici proportionnelle à la température absolue.) Les courbes que l'on obtient ainsi sont qualitativement en accord avec les prévisions. Toutefois, la pente des résistivités

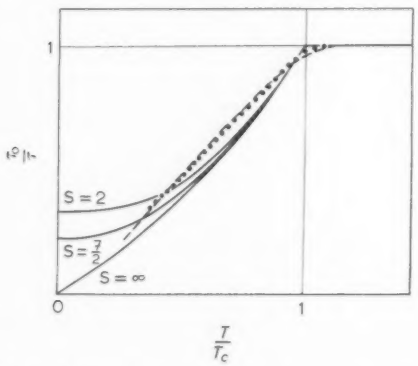


FIG. 2. Résistivité par désordre de spin dans l'approximation élastique. Trait continu : valeurs déduites de (2.13) en calculant l'aimantation spontanée par une approximation de champ moléculaire. - - - : valeurs expérimentales pour Gd. ..... : valeurs expérimentales pour AuMn (après soustraction approchée de la résistivité due aux phonons).

théoriques en dessous de T_c est *trop forte*; (elle varie de 1,66 pour $S = \infty$ à 1,58 pour $S = 5/2$.

Expérimentalement, la section efficace incohérente déduite de ces comparaisons, est d'environ 6,5 Å² pour le Gadolinium, et de 4 Å² pour le manganèse dans Au Mn. Dans l'approximation de BORN, ceci correspond à des valeurs anormalement élevées (de 14 à 18 eV . Å³) de la constante de couplage G.

En fait, la diffusion est probablement contrôlée par un effet de résonance avec les niveaux virtuels d ou f de l'ion magnétique.[9] La méthode simple que nous avons adoptée ne peut pas rendre compte de tels effets.

Notons à ce propos que la courbe expérimentale de résistivité pour Au^3Mn[2] permet de comparer les effets du désordre de spin qui apparait vers 150°F, et du désordre dans la répartition spatiale des atomes Mn qui apparait vers 900° K. Le désordre de spin crée ici une diffusion 5 *fois plus forte* que le désordre de position.

5. EFFETS D'ORDRE A COURTE DISTANCE

Prenons pour exemple le cas d'un ferromagnétique au-dessus du point de transition. Fixons un spin suivant une certaine direction au site O. Les autres spins prennent une orientation moyenne

$$\bar{\boldsymbol{S}}_R = \boldsymbol{S}_0 \gamma_R \qquad (5.1)$$

Ecrivons que le spin $\boldsymbol{S}_R$ est en équilibre sous l'action du champ moléculaire dû à ses voisins (R'). Nous trouvons :[10]

$$\bar{\boldsymbol{S}}_R = \frac{2JS(S+1)}{k_B T} \sum_{R'} \bar{\boldsymbol{S}}_{R'} \qquad (5.2)$$

ou encore, comme

$$k_B T_c \cong \tfrac{2}{3} JvS(S+1)$$

$$\gamma_R = \frac{1}{V} \frac{T_c}{T} \sum_{R'} \gamma_{R'}. \qquad (5.3)$$

A partir de cette équation aux différences finies, on peut obtenir directement la quantité qui intervient dans la section efficace :

$$\sum_R \gamma_R \exp(i \boldsymbol{\varkappa} . \boldsymbol{R}) = \left(\sum_R \gamma_R \right) \frac{L(0)}{L(\boldsymbol{\varkappa})}. \qquad (5.4)$$

* Pour le gadolinium, on peut utiliser dans (4.3) la valeur expérimentale[13] de $\langle \boldsymbol{S} \rangle_T$. Il n'y a d'écart notable avec le résultat de l'approximation de champ moléculaire que dans la région de basses températures, où de toute façon le traitement n'est pas applicable.

où :

$$L(\boldsymbol{\varkappa}) = 1 - \frac{T_c}{T}\frac{1}{V}\sum_{R'} \exp(i\boldsymbol{\varkappa}.\boldsymbol{R}') \qquad (5.5)$$

($\boldsymbol{R}'$ = voisin de O)

La quantité $\sum_R \gamma_R$ peut être déterminée par l'étude des fluctuations macroscopiques de l'aimantation.[11] A l'approximation du champ moléculaire, on a finalement

$$\sum_R \gamma_R \exp(i\boldsymbol{\varkappa}.\boldsymbol{R}) = \frac{1}{L(\boldsymbol{\varkappa})}. \qquad (5.6)$$

Dans la pratique, il sera d'ailleurs suffisant de remplacer $L(\boldsymbol{\varkappa})$ par sa moyenne sur toutes les orientations due vecteur $\boldsymbol{\varkappa}$. On prendra donc :

$$\sum_R \gamma_R \exp(i\boldsymbol{\varkappa}.\boldsymbol{R}) = \frac{1}{1 - \dfrac{T_c}{T}\dfrac{\sin(\kappa\alpha)}{\kappa\alpha}}. \qquad (5.7)$$

$$\frac{\tau_0}{\tau} = \frac{1}{4}\int_0^2 \frac{x^3 dx}{1 - \dfrac{T_c}{T}\dfrac{\sin(k_0 dx)}{k_0 dx}} \qquad (5.8)$$

où l'on a posé $x^2 = 2(1-\cos\theta)$.

Pour les autres situations ($T < T_c$, antiferromagnétiques, etc.) on peut développer le calcul des corrélations par des méthodes analogues.[10][12]

Les résultats numériques sont résumés sur les figures (3) et (4). La conclusion essentielle est la suivante : pour $k_0 d \gtrsim \pi$ l'effet des corrélations est *faible*. Au contraire, pour $k_0 d \ll \pi$ on attend un fort accroissement de la résistivité au voisinage du point de CURIE. Ceci est essentiellement lié aux phénomènes d'opalescence critique,[11][12] qui apparaissent aux petits angles et aux grandes longueurs d'onde. Dans ce domaine, l'approximation de champ moléculaire doit d'ailleurs être améliorée, soit par la méthode de BETHE et PEIERLS (ELLIOTT-MARSHALL, non publié), soit par l'usage des susceptibilités expérimentales.

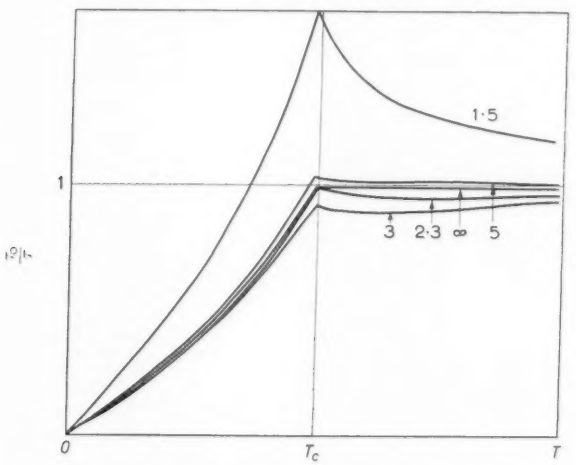


FIG. 3. Effet de l'ordre à courte distance sur la résistivité d'un ferromagnétique. Approximation élastique; traitement de champ moléculaire; spin classique. Les courbes ont été tracées pour quelques valeurs du paramètre $k_0 d$ (marquées sur le graphique).

En reportant ce résultat dans (2.7), on obtient une valeur approchée (à cause de l'approximation de champ moléculaire) mais qui tient compte des corrélations à toute distance. Le temps de relaxation est donné par :

Pour un réseau compact, on a approximativement : $k_0 d = 3{,}5\xi^{\frac{1}{3}}$. Dans les cas pratiques, la condition $k_0 d \gtrsim \pi$ sera donc satisfaite en général.

Enfin, il parait difficile, pour les cas que nous avons envisagés, d'extraire les effets de corrélation

des courbes expérimentales car ils ne présentent pas de fortes singularités au point de transition, et leur décroissance au-dessus du point de transition est lente (comme $1/T$ aux températures élevées). Pour Gd, l'accord avec l'expérience n'est pas sensiblement amélioré par l'introduction des corrélations.

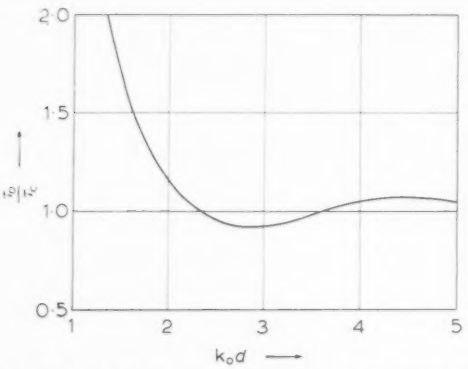


FIG. 4. Résistivité au point de transition divisée par la valeur limite de haute température, en fonction du paramètre $k_0 d$.

6. EFFET MAGNÉTOVIBRATIONNEL

Cet effet (introduit par ELLIOTT et LOWDE[14] à propos des neutrons) est une diffusion dûe à l'interaction (1.1) mais avec émission ou absorption de phonons (le réseau des spins restant lui, non perturbé). Nous allons montrer qu'il est faible. La section efficace correspondante vaut (dans l'approximation à un phonon, et pour un cristal ou il y aurait une seule fréquence $\omega_{\boldsymbol{\varkappa}}$ de vibration pour tous les modes associés au vecteur d'onde $\boldsymbol{\varkappa}$)

$$\frac{d\sigma_1}{d\Omega} = \frac{\sigma_0}{4\pi} \frac{\langle \boldsymbol{S} \rangle_T^2}{S(S+1)} \frac{\hbar\kappa^2}{M\omega_\kappa} (n_\kappa + \tfrac{1}{2}) \qquad (6.1)$$

M est la masse d'un atome du réseau, n_κ le nombre moyen de quanta de l'oscillateur. En ordre de grandeur, pour des températures supérieures à la température de DEBYE, on peut écrire

$$\frac{\sigma_1}{\sigma_0} \cong \frac{\langle \boldsymbol{S} \rangle_T^2}{S(S+1)} \frac{k_B T}{MC^2} \qquad (6.2)$$

où C est la vitesse de son. On en déduit que σ_1 est nul au dessus de T_c et toujours *beaucoup plus petit* (par un facteur de l'ordre de 100) que σ_0; cependant la variation de σ_1 est rapide au voisinage inférieur de T_c. Un tel effet tend à améliorer l'accord avec les courbes expérimentales, mais il est trop faible et son évaluation précise est exclue.

7. REMARQUES GÉNÉRALES

(1) *Region des basses températures*

Pour $T \to 0$ formule (4.3) déduite de l'approximation élastique prévoit :

$$\frac{\tau_0}{\tau} = \frac{1}{S+1}.$$

Ceci correspond au désordre résiduel, dû aux oscillations de point O des ondes de spin. En fait, comme nous l'avons vu à propos des transferts d'énergie, l'emploi de (4.3) suppose que les électrons sont capables d'échanger avec les ions des énergies de l'ordre de $[\langle (\hbar\omega)^2 \rangle]^{\frac{1}{2}}$. Ici ce n'est plus le cas, et l'on doit résoudre en détail une équation de transport. (Le calcul correspondant sera publié à part.) Le fait essential est que dans ce domaine, la résistivité est *très sensible aux détails du modèle physique*.

Par exemple, pour un ferromagnétique, si les ondes de spin ne sont pas entrainées par le courant, et si les électrons ne sont *pas* polarisés, la résistivité est proportionnelle à T^2 aux basses températures : dans un choc, la composante longitudinale du transfert d'impulsion est proportionnelle à κ^2. L'angle solide qui correspond aux déflexions moyennes est aussi proportionnel à κ^2. La résistivité est proportionnelle à κ^4. Pour les ondes de spin avec lesquelles le transfert d'énergie est possible, κ est proportionnel à $T^{\frac{1}{2}}$ (Pour des phonons, la probabilité de transition contient en plus un facteur κ; κ est proportionnel à T, et ρ est donc proportionnel à T^5.) Par contre, si les électrons *sont polarisés*, on s'attend à ce que la résistivité par désordre de spin varie comme $\exp(-E/k_B T)$ aux basses températures. En effet, les deux demi-bandes correspondent à des nombres d'électrons différents ξ_+ et ξ_-, et les vecteurs d'onde au bord de la zone de Fermi k_{0+} et k_{0-} sont différents. Le retournement d'un spin d'un électron de conduction exige l'absorption ou l'émission d'une onde de spin dont le vecteur d'onde est au minimum $|k_{0+} - k_{0-}|$ en longueur. L'émission d'une onde de spin est limitée par le peu d'énergie dont dispose un électron au voisinage du niveau de Fermi; l'absorption l'est par le petit nombre d'ondes de spin présentes. L'énergie d'activation correspondante est

approximativement égale à l'énergie d'une onde de spin dont le vecteur d'onde aurait la longueur $k_{0+} - k_{0-}$.

Ceci donne :

$$E \sim \frac{2k_B T_c}{9(S+1)}(k_0 d)^2 \left(\frac{\xi_+ - \xi_-}{\xi_+ + \xi_-}\right)^2$$

(2) *Critique des approximations utilisées*

L'usage d'un potentiel en δ et de l'approximation de Born décrit mal la diffusion par un spin isolé et imparfaitement les effets d'interférence entre diffuseurs différents. Comme pour les alliages à fortes concentrations, il est difficile d'améliorer cette situation sans reprendre à la base l'étude de la propagation dans un milieu désordonné.

On peut se demander quelles seraient les modifications apportées si l'on adoptait un potentiel $V(R)$ plus réaliste et si l'on traitait rigoureusement la diffusion par *un* centre. Les effets de corrélation sont modifiés; ils ne peuvent pas être calculés exactement, mais ils restent petits. Notre traitement implique d'ailleurs plusieurs autres hypothèses également essentielles : les corrélations entre spins ioniques et spins des électrons de conduction sont négligées. La bande de conduction est supposée *non déformée* par l'ordre à longue distance des ions magnétiques. En particulier, dans un antiferromagnétique, il y a en fait une subdivision de la zone de BRILLOUIN dont nous n'avons pas tenu compte. L'analogie des courbes expérimentales de résistivité pour AuMn et Au_3Mn, où la configuration géométrique n'est pourtant pas la même, tend à prouver que, dans ces alliages, il n'y a pas de distortion sensible au voisinage du niveau de Fermi.

REFERENCES

1. LEVGOLD S., SPEDDING F. H., BARSON F., et ELLIOTT J. F. *Rev. Mod. Phys.* **25,** 129 (1953). Voir aussi l'article de revue dans *Progress in Nuclear Energy*. Vol. V. Pergamon Press, London.
2. GIANSOLDATI A. et LINDE J. O. *J. Phys.* **16,** 341 (1955).
3. HOFFMANN J. A., PASKIN A., TAUER K. J. et WEISS R. J. *J. Phys. Chem. Solids* **1,** 45 (1956).
4. FRIEDEL J. *J. Phys.* **16,** 829 (1955).
5. MOTT N. F. *Proc. Roy. Soc.* A **153,** 699 (1956).
6. VAN VLECK J. H. *Phys. Rev.* **55,** 924 (1939).
7. DE GENNES P. G., *C. R. Acad. Sci., Paris* **244,** 755 (1957).
8. NORDHEIM L. *Ann. Phys.* **9,** 607 (1931).
9. FRIEDEL J. *Canad. J. Phys.* **34,** 1190 (1956).
10. DE GENNES P. G. et HERPIN A. *C. R. Acad. Sci., Paris* **243,** 1611 (1956).
11. VAN HOVE L. *Phys. Rev.* **95,** 1374 (1954).
12. DE GENNES P. G. Rapport SPM Saclay n°200.
13. ELLIOTT J. F., LEVGOLD S. et SPEDDING F. H. *Phys. Rev.* **91,** 28 (1953).
14. ELLIOTT R. J. et LOWDE R. D. *Proc. Roy. Soc.* A **230,** 46 (1955).
15. PARMENTER R. H. *Phys. Rev.* **97,** 587 (1955).
16. KORRINGA J. *Canad. J. Phys.* **34,** 1290 (1956).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 78–90.

ADSORPTION AND CHARGE TRANSFER ON SEMICONDUCTOR SURFACES

H. J. KRUSEMEYER and D. G. THOMAS

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

(*Received* 3 *June* 1957)

Abstract—The concentration of ions and neutral atoms adsorbed on a semiconductor surface, as well as its surface potential, have been treated as a function of the pressure of surrounding gases for coverages up to one monolayer, considering the occupation of sites by both ionized and unionized atoms. Pure gases and a mixture of two gases, one giving rise to positive, the other to negative adions, are treated. As an example, numerical results are obtained for semiconductors such as ZnO with a large forbidden gap at temperatures below the intrinsic range. For a gas with pressure P_d and with r atoms per molecule which give rise to positive adions on an n-type semiconductor, the adion density is proportional to $P_d^{1/r}$ for very low coverages. Following a transition region, the coverage becomes proportional to $P_d^{1/3r}$ until degeneracy of the electron gas near the surface further impedes coverage increase. For gases giving rise to negative adions, the behavior at low coverage is the same as for gases giving positive adions. This region and a transition range are followed by a region where the ion coverage is practically independent of pressure until an inversion layer is formed. From this point on, the behavior is again similar to that at higher coverage for gases which give positive ions. The evaluation of the energy of electronic levels formed by the gas atoms, the total number of adsorption sites, and the influence of fast states are discussed.

1. INTRODUCTION

IT is well known[1,2] that atoms adsorbed on the surface of a semiconductor can affect both its electrical conductivity and its work function. One of the models proposed is that in which adsorbed atoms give rise to localized levels for electrons at the surface.[1–5] These levels can be filled by current carriers, which are thus immobilized, or they can be emptied and so provide extra carriers. The field of these adsorbed ions sets up a space-charge layer near the surface on which it terminates. The voltage drop across this space-charge layer represents a change in the work function of the surface and changes in turn the energy of the surface levels with respect to the Fermi level. HAUFFE and ENGELL[5] have used this model to derive the dependence of adsorption on gas pressure. These authors consider for oxygen the reaction $\frac{1}{2}O_{2\,gas} \rightleftarrows O^{=}_{ads} - 2\epsilon$ (electron) with an equilibrium constant

$$K_0 = \frac{[O^{=}_{ads}]}{n_0^{-2} P_{O_2}^{\frac{1}{2}}}.$$

Here n_0^- is the concentration of conduction electrons at the surface and P_{O_2} the pressure of the oxygen gas. Both $[O^{=}_{ads}]$ and n_0^- can be expressed in terms of the potential difference between the surface and interior. Consequently n_0^- can be expressed in terms of $[O^{=}_{ads}]$ without involving the potential drop across the sample. To obtain this relation between $[O^{=}_{ads}]$ and n_0^-, HAUFFE and ENGELL assumed certain artificial charge distributions in the space-charge layer, the limits of validity of which are not determined, so that the resulting equations are applicable only over certain unknown ranges. Furthermore saturation of the ion adsorption because of a limited number of sites has not been considered. Such saturation effects were observed by THOMAS and LANDER.[6]

We, on the other hand, choose to define an equilibrium constant involving the reaction of the gas with the neutral species $\frac{1}{2}O_2 \rightleftarrows O_{ads}$ with the equilibrium constant

$$K_1 = \frac{[O_{ads}]}{P_{O_2}^{\frac{1}{2}} \left(1 - \frac{\sigma}{\sigma_0}\right)^{\frac{1}{4}}},$$

σ_0 being the total number of sites at the surface and σ the number of sites occupied by both atoms and ions. The subsequent ionization of the adsorbed atoms is expressed in terms of the Fermi statistics. It was found that the HAUFFE and ENGELL treatment holds for gases giving donor-type surface states on n-type and acceptor states on p-type material if the voltage drop across the space-charge layer is larger than $\approx 2kT/e$ and as long as the assembly of conduction electrons or holes is non-degenerate and no saturation effects come into play. The lower limit of applicability for gases giving acceptor-type states on n-type and donor states on p-type material was $\approx 4kT/e$. As a given semiconductor may adsorb both donor- and acceptor-type gases and as no gas is quite pure, the treatment below is for a mixture of two gases, one giving rise to donor-type the other to acceptor-type surface states.

Observations on germanium and silicon, for which the interior of the semiconductor was screened more effectively from the surface field than was anticipated, led to the assumption of fast states.[1,7] It seems likely that Shockley-type surface states will play a less important role in polar or semi-polar crystals like zinc oxide than in pure covalent semiconductors like germanium. However, the fast states may arise for instance too from chemical impurities not in equilibrium with the surrounding gas atmosphere. For the argument presented here, it is not important whether these states are more or less tightly coupled to the conduction electrons or holes of the semiconductor than the states caused by adsorbed gas atoms. They will, however, still be called fast states. The important property of these states made use of below is that their total number is assumed to be constant, in contrast to the number of states from gas atoms, which is dependent on gas pressure.

While the treatment to be presented will be general, we think it likely that the effects to be discussed will be more prominent with ionic than with covalent semiconductors. Thus, if an oxygen or a hydrogen atom is to be chemisorbed on a clean germanium surface (no oxide layer) and if each germanium atom in the surface has one odd electron, then a Ge–O or a Ge–H bond can be formed without there being any surplus or deficiency of electrons. If, however, the surface atoms have completed electronic shells, as may be the case for ionic oxides, then the formation of covalent bonds or the extension of the ionic lattice at the surface involves the transfer of conduction electrons to localised levels. Thus, zinc may be adsorbed on zinc oxide by the reaction

$$\mathrm{Zn} \rightarrow \mathrm{Zn}^{+} + \epsilon$$

with the liberation of one or perhaps two electrons/atom, or oxygen may be adsorbed as O^- or $O^=$ with the removal of one or two electrons/atom from the crystal. Partly because of these conditions, ZnO in an atmosphere of zinc and oxygen has been chosen as an example to which the general treatment is applied and for which numerical results are presented.

2. OUTLINE OF TREATMENT

We will consider a semiconductor of semi-infinite extent bounded by a flat infinite plane on which gas may be adsorbed. The semiconductor, which may be p-type, n-type, or intrinsic, is in contact with a mixture of two gases. The molecules of one gas, with a partial pressure P_d, consist of r identical atoms. If such a molecule approaches the surface, it may split up according to the model of LENNARD-JONES into r atoms, each of which gives rise to a donor-type level at the surface. The molecules of the second gas, with pressure P_a, consist of s identical atoms, each of which may cause an acceptor-type level at the surface. The adatoms giving rise to these levels, whether ionized or not, are supposed to be tightly bound to the surface, almost as though they were an extension of the lattice. For instance, in the unionized condition of a donor atom, the orbit of the odd electron is thought to stretch over many atomic

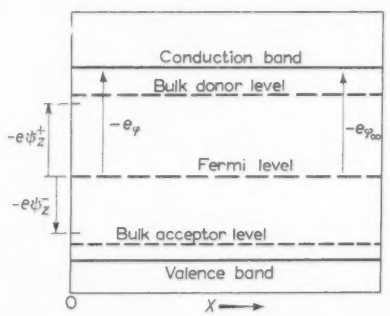


FIG. 1. Electron energy diagram for a neutral surface.

diameters. The electron-energy diagram in Fig. 1 shows, for the special case of a neutral surface, these two surface levels respectively with energies $-e\psi_z^+$ and $-e\psi_z^-$. As we assume thermodynamic equilibrium, zero energy has been chosen to coincide with the Fermi level; e is the magnitude of the electron charge. The positions of the levels in the forbidden gap have been selected arbitrarily, but are assumed to be fixed with respect to the band structure of the semiconductor at the surface and to move with it as shown in Fig. 2, if bending of

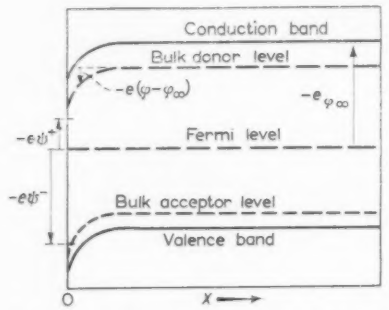


FIG. 2. Electron energy diagram for a positively charged surface.

the bands is caused by an electric field penetrating the semiconductor. In this figure the partial pressures of the components of the gas phase around the semiconductor are such that the surface is positively charged. The electric field set up by the surface charge in both slow and fast states will cause a rearrangement of the charge in the interior and so lead to a suppression of the electric field in the semiconductor. As the density of carriers in a semiconductor is usually low, the field of the surface charge will be able to penetrate far enough to set up a space-charge layer and cause a change in the potential of the surface ϕ_0. This change in surface potential is assumed to be solely responsible for the change of total charge in fast states.

The energy difference between the edge of the conduction band and the Fermi level is $-e\phi$, where ϕ is a function of position and is defined as the electrostatic potential. In the field-free interior its value is denoted by ϕ_∞. The energy differences between the donor and acceptor surface states due to adsorbed gases and the Fermi level are indicated as $-e\psi^+$ and $-e\psi^-$. ψ_z^+ and ψ_z^- in Fig. 1 are, of course, the particular values for ψ^+ and ψ^- if the voltage drop across the space-charge layer is not taken into account.

If the density of the surface levels becomes appreciable, they may interact and thus spread out along the surface to form a narrow band. Before this state is reached, the energy of a level may already depend on peculiarities of the local environment. It will be assumed here, however, that the levels of donor-type adatoms have all the same energy. The same assumption is made for the levels of acceptor-type adatoms.

When simultaneous adsorption is considered of two gases of different type, as in the case of oxygen and zinc adsorbed on zinc oxide, the gases may react together as well as the ions at the surface. Complications may thus arise, depending on the rates of the different processes involved. Here it will be assumed that the complicating reactions are proceeding slowly enough to be neglected.

By the adsorption of these gases, carriers will be added to or removed from the semiconductor. Since electrical neutrality of the whole semiconductor is maintained, the surface charge must equal the charge added to or removed from the semiconductor, and so is closely correlated with the change in conductivity. However, even for the simple case where the change in the number of surface charges equals the change in the number of carriers (no trapping), the relation between ion adsorption and conductivity is complicated by a dependence of the carrier mobility on the shape of the space-charge layer which has been treated by SCHRIEFFER.[8]

Another consequence of the interaction of surface atoms is that conduction may also take place in the band of surface levels so formed.

The density of neutral donor and acceptor species adsorbed at the surface will be denoted by p and m, and the density of donor and acceptor ions in equilibrium with these neutral species by p^+ and m^-. The values for p^+ and m^- for certain values of p and m are, as we will see, partly determined by the voltage drop across the space-charge layer. p_z^+ and m_z^- are the values obtained for p^+ and m^-, keeping p and m constant, if the influence of the space-charge on the surface potential is not taken into account. This case is depicted in Fig. 1.

The calculation of the number of charges at the surface will, for reasons of convenience, be done in two distinct steps. In the first step the reaction

VOL. 4 1958

between the molecules or atoms in the gas phase with the neutral species at the surface will be considered. From the densities of the neutral species, p_z^+ and m_z^- will be derived. Thus we obtain a relation between p_z^+ and m_z^- on the one hand and the partial gas pressures on the other. In the second step $p^+ - m^-$, the net surface density of ions, taking the voltage drop across the space-charge layer into account, will be obtained from p_z^+ and m_z^-. This involves the calculation of the barrier height of the space-charge layer as a function of the density of the net surface charge both in adions and in fast states. This procedure has some advantage in that the properties relating to the interaction of the gas molecules and atoms with the semiconductor surface appear only in the first step of the calculation, the second step being determined solely by the properties of the semiconductor. We will assume that each of these surface atoms ionizes only once. Double ionization is in principle possible, but it would not change the following treatment materially. M.K.S. units will be used throughout. A list of symbols appears at the end.

3. FERMI STATISTICS OF THE SURFACE LEVELS

If equilibrium is assumed between the gas phase and the adsorbed species, the Langmuir adsorption model[9] gives for the density of neutral donor and acceptor adatoms

$$p = K_1\left(1-\frac{\sigma}{\sigma_0}\right)_d^{1/r} P_d^{1/r}$$

$$\text{and} \quad m = K_2\left(1-\frac{\sigma}{\sigma_0}\right)_a^{1/s} P_a^{1/s}. \tag{1}$$

K_1 and K_2 are equilibrium constants and are assumed independent of coverage, σ/σ_0 is the fraction of adsorption sites occupied by neutral and ionized donors or neutral and ionized acceptors. Equation (1) is in effect true only if $r = s = 1$. If $r \neq 1$, a gas molecule hitting the surface would probably need r adjoining sites to accomodate its r atoms. This is not taken into account in deriving equation (1). However, as will be shown later, the saturation region, i.e., the region in which $1-(\sigma/\sigma_0) \neq 1$, is coveragewise very narrow, and accordingly it is more important to know at what ion coverage the saturation takes place than the actual shape of the adsorption isotherm in the saturation region. Furthermore, of particular interest here are gases with one or two atoms per molecule.

Fermi statistics give for the number of electrons in surface states of adatoms

$$p = \frac{p+p^+}{1+D\exp[-e\psi^+/kT]} \qquad \text{and}$$

$$m^- = \frac{m+m^-}{1+A\exp[-e\psi^-/kT]}$$

or

$$p = D^{-1}p^+\exp[e\psi^+/kT];$$
$$m = Am^-\exp[-e\psi^-/kT]. \tag{2a}$$

The constant D in the Fermi function is the ratio of the numbers of electronic states of a donor in the neutral and ionized condition respectively. A is the ratio of the number of hole states for an ionized and neutral acceptor. If the voltage drop across the space-charge layer is not taken into account,

$$p = D^{-1}p_z^+\exp[e\psi_z^+/kT];$$
$$m = Am_z^-\exp[-e\psi_z^-/kT]. \tag{2b}$$

Combining (1) and (2b) gives

$$p_z^+ = DK_1\exp[-e\psi_z^+/kT]\left(1-\frac{\sigma}{\sigma_0}\right)_d^{1/r} P_d^{1/r} \tag{3}$$

and

$$m_z^- = A^{-1}K_2\exp[+e\psi_z^-/kT]\left(1-\frac{\sigma}{\sigma_0}\right)_a^{1/s} P_a^{1/s}.$$

It only remains to express p^+ and m^- in p_z^+ and m_z^-. It can be easily seen from Figs. 1 and 2 that

$$-e\psi^\pm = -e(\psi_z{}^\pm + \phi_0 - \phi_\infty) \tag{4}$$

which gives, in combination with (2a) and (2b),

$$\frac{p_z^+}{p^+} = \frac{m^-}{m_z^-} = \exp[e(\phi_0-\phi_\infty)/kT]. \tag{5}$$

The height of the space-charge barrier will be constant if the partial pressures P_d and P_a are changed in such a way that the net charge of adsorbed gas ions is zero, i.e. $p^+ = m^-$ or as equation (5) shows p_z^+/m_z^- is constant. Thus there is,

according to equation (3), outside the region for which $1-(\sigma/\sigma_0) \neq 1$, a critical value for $P_d^{1/r}/P_a^{1/s}$, independent of P_d or P_a, for which the gas ions do not contribute to the net surface charge. Furthermore equations (3) and (5) show that, in case the conductivity of a semiconductor is constant for all partial pressures which obey a relation $P_d^{1/r}/P_a^{1/s}$ = constant (any value), the gas ions do not contribute to the net surface charge for these partial pressures.

The barrier height $\phi_0-\phi_\infty$ of the space-charge layer in equation (5) will be obtained in the next section by integrating Poisson's equation once across the space-charge layer, using the field caused by the net charge density both in fast states and in adions as one boundary condition.

We will assume both donor- and acceptor-type fast states which are not shown in Figs. 1 and 2. $q(\zeta_z{}^+)$ is the density of fast states, both charged and neutral, which have an energy $-e\zeta_z{}^+$ with respect to the Fermi level if the voltage drop across the space-charge layer is not taken into account. $q^+(\zeta_z{}^+)$ is the density of the charged fraction of these donors.
Fermi statistics gives

$$q(\zeta_z{}^+)-q^+(\zeta_z{}^+) = \frac{q(\zeta_z{}^+)}{1+D(\zeta_z{}^+)\exp[-e\zeta^+/kT]}$$

and as these states too move with the bending of the bands

$$-e\zeta^+ = -e\zeta_z{}^+-e\phi_0+e\phi_\infty.$$

Thus with the use of equation (5)

$$\sum q^+(\zeta_z{}^+) = \sum_{\zeta_z{}^+} \frac{q(\zeta_z{}^+)}{1+D^{-1}(\zeta_z{}^+)\dfrac{p_z{}^+}{p^+}\exp[e\zeta_z{}^+/kT]}. \tag{6a}$$

The summation takes the possibility of different energy levels into account. Similarly we obtain for the density of the surface charge in fast surface acceptors

$$\sum q^-(\zeta_z{}^-) = \sum_{\zeta_z{}^-} \frac{q(\zeta_z{}^-)}{1+A(\zeta_z{}^-)\dfrac{p^+}{p_z{}^+}\exp[-e\zeta_z{}^-/kT]} \tag{6b}$$

Both $\zeta_z{}^+$ and $\zeta_z{}^-$ are, of course, material constants of the semi-conductor independent of the surrounding gases.

4. BARRIER HEIGHT OF THE SPACE-CHARGE LAYER

$\rho^+(\phi)$ and $\rho^-(\phi)$ represent respectively the charge density of positive and negative impurities in the semiconductor as a function of electrostatic potential. Thus Poisson's equation can be written as

$$\frac{\partial^2\phi}{\partial x^2} = -\frac{[\rho^+(\phi)+\rho^-(\phi)]}{\epsilon\epsilon_0} \tag{7}$$

ϵ and ϵ_0 are, respectively, the dielectric constant of the semiconductor and the permittivity of a vacuum ($8{\cdot}85\times10^{-12}$ F/m).

Integrating (7) gives

$$\left(\frac{\partial\phi}{\partial x}\right)_0^2 = \frac{2}{\epsilon\epsilon_0}\int_{\phi_0}^{\phi_\infty}[\rho^+(\phi)+\rho^-(\phi)]\,d\phi.$$

Only the surface field appears in the equation because the field in the interior goes to zero. As the charge on the surface just balances that in the semiconductor, by using Gauss's theorem the left-hand side can be expressed in terms of the surface-charge density. The right-hand side can be written as a function of $\phi_0-\phi_\infty$ and so, according to equation (5), as a function of $p_z{}^+/p^+$.

Thus

$$p^+-\frac{m_z{}^-p_z{}^+}{p^+} = -\sum q^+(\zeta_z{}^+)+ \\ +\sum q^-(\zeta_z{}^-)+F\left(\frac{p_z{}^+}{p^+}\right) \tag{8}$$

$$F\left(\frac{p_z{}^+}{p^+}\right) = \pm\left[\frac{2\epsilon\epsilon_0}{e^2}\times \int_{\phi_0}^{\phi_\infty}\{\rho^+(\phi)+\rho^-(\phi)\}\,d\phi\right]^{\frac{1}{2}}.$$

The plus sign is to be used for positive surface-charge only.

$$\int_{\phi_0}^{\phi_\infty}\{\rho^+(\phi)+\rho^-(\phi)\}\,d\phi$$

has been obtained by various authors for a number of special cases.[7,10] GARRETT and BRATTAIN obtained this relation without assuming equilibrium between the holes and electrons in the space-charge layer, although the electrons and the holes were each in equilibrium with themselves throughout this region; that is, the quasi-Fermi levels for holes and electrons are horizontal but separate in the space-charge layer. GARRETT and BRATTAIN also made the assumptions of immobile and homogeneously distributed donors and acceptors which were ionized at all points, and that Boltzmann statistics were applicable to the free holes and electrons.

Assuming that thermddynamic equilibrium exists, but making no restrictions as to the degree of ionization of the donor and acceptor levels, we obtain

$$F\left(\frac{p_z^+}{p^+}\right) = \pm\frac{(2kT\epsilon\epsilon_0)^{\frac{1}{2}}}{e}\times$$

$$\left[n_\infty^-\left(-1+\frac{p_z^+}{p^+}\right)+p_\infty\left(-1+\frac{p^+}{p_z^+}\right)+\right.$$

$$+\sum_i N_{di}\left\{\ln\left(\alpha_i-1+\frac{p^+}{p_z^+}\right)-\ln\alpha_i\right\}+$$

$$\left.+\sum_i N_{ai}\left\{\ln\left(\beta_i-1+\frac{p_z^+}{p^+}\right)-\ln\beta_i\right\}\right]^{\frac{1}{2}} \qquad (9)$$

where n_∞^- and p_∞ are respectively the electron and hole densities in the field-free interior of the semiconductor, N_{di} is the density of donors of which a fraction $1/\alpha_i$ is ionized in the field-free interior, and N_{ai} and $1/\beta_i$ are these quantities for acceptors. The Σ signs take into account the presence of more than one donor or acceptor level. Expression (9) can easily be derived by substituting the contribution of the different species to the space charge, in equation (8), and by performing the simple integrations and eliminating $\phi_0-\phi_\infty$ with the aid of equation (5).

Non-degeneracy of the conduction electrons and holes, as well as a homogeneous distribution of immobile donors and acceptors, has been assumed. Expressions for $F(p_z^+/p^+)$ for enrichment or inversion layers with degeneracy of the conduction electrons or holes are given in the Appendix.

The use of the one-dimensional form of Poisson's equation (7) presumed the electric field to be homogeneous right up to the surface and thus the surface-charge to be smeared out. As the surface charge can be more accurately represented by point charges, it may be expected that the time average of the screening charge will be distributed approximately hemispherically around these point charges. The one-dimensional approach is justified if the distance between the point charges on the surface is small compared to the radius of a hemisphere of space charge which will screen out the electric field of one of the point charges. In case the opposite is true, the surface charges will be screened from each other, the adsorption of ions becomes to a lesser degree a co-operative phenomenon, and the barrier height of the space-charge layer for a certain surface-charge density will be less than the value given by equation (8). This situation is most likely to develop for relatively low coverages with positive ions and a large density of the conduction electrons in the field-free interior of the semiconductor.

Equations (2b), (3), (5), (6a), (6b), and (8) give the relation between p^+ and m^- on the one hand and P_d and P_a on the other.

As equation (8) cannot, in general, be written explicitly in p^+, numerical results can be obtained as follows. Assume a value for m_z^- and for p_z^+/p^+. Using equations (8), (6a), and (6b), p^+ is obtained and thus p_z^+. These values substituted in equation (5) give m^-. Subsequently equation (2b) yields p and m. Finally, p_z^+, m_z^-, p^+, p, m^-, and m in equation (3) give P_d and P_a.

In case $P_d = 0$, it is necessary to write m^-/m_z^- for p_z^+/p^+ in equations (6a), (6b), and (8). Equation (3) shows that p_z^+ and m_z^- are proportional to $P_d^{1/r}$ and $P_a^{1/s}$ except for the region where σ/σ_0 becomes of the order of 1. Thus for the same semiconductor but different pure gases giving rise to positive adions and having the same number of atoms per molecule, the ion-adsorption isotherms outside the saturation region differ only by a constant factor. The same holds for gases producing negative surface ions.

5. NUMERICAL RESULTS FOR ZnO

For an n-type semiconductor with only one kind of donor which is completely ionized in the field-free interior, and with a large enough forbidden gap so that intrinsic conduction can be neglected in the field-free region, equations (8) and (9) can be written as:

$$p^+ - \frac{m_z^- p_z^+}{p^+} = -\sum q^+(\zeta_z^+) + \sum q^-(\zeta_z^-) \pm \pm\beta^{\frac{1}{2}}\left[-1 + \frac{p_\infty p^+}{n_\infty^- p_z^+} + \frac{p_z^+}{p^+} + \ln\frac{p^+}{p_z^+}\right]^{\frac{1}{2}}$$

$$\text{or } \frac{p_z^+ m_z^-}{m^-} - m^- = -\sum q^+(\zeta_z^+) + \sum q^-(\zeta_z^-) \pm\beta^{\frac{1}{2}}\left[-1 + \frac{p_\infty}{n_\infty^-}\frac{m_z^-}{m^-} + \frac{m^-}{m_z^-} + \ln\frac{m_z^-}{m^-}\right]^{\frac{1}{2}} \quad (10)$$

where $\beta = 2kT\epsilon\epsilon_0 n_\infty^-/e^2$.

forbidden gap are self-consistent. The numbers against points in Fig. 3 are the barrier height of the space-charge layer $\phi_0 - \phi_\infty$ in units of kT/e as obtained from equation (5). At 600°K $kT/e = 0{\cdot}052$. For pure zinc vapor p^+ is proportional to p_z^+ for small values of p_z^+, as the change in surface potential is too small to influence the occupation of the surface levels. The straight section of the curve for larger values of p_z^+ corresponds with $p^+ = \beta^{\frac{1}{2}} p_z^{+\frac{1}{2}}$. As the density of the conduction electrons near the surface becomes large enough to cause degeneracy, Fermi–Dirac statistics should be used for calculating $\rho^-(\phi)$ in equation (8). The

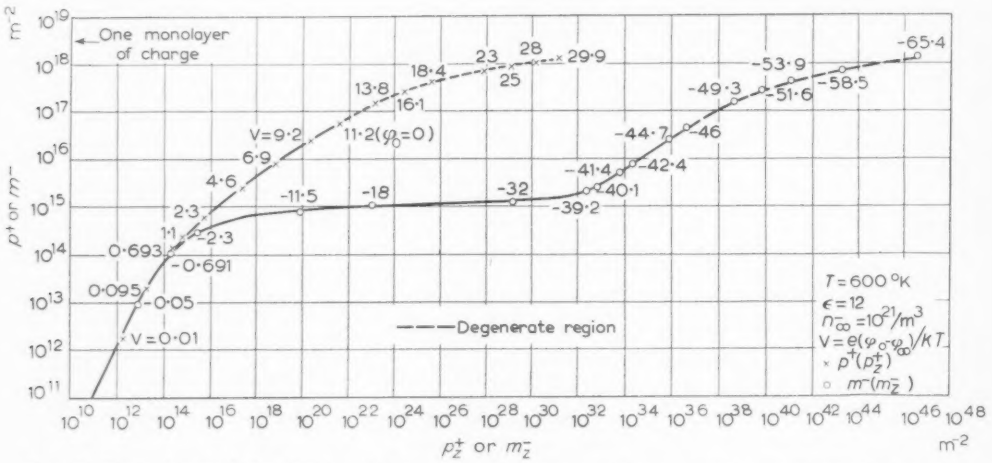


FIG. 3. Dependence of the density of the net surface charge in electronic units on the quantities p_z^+ and m_z^- respectively for pure zinc and pure oxygen vapor.

As the gap width in ZnO is 3eV,* the results shown below may apply to ZnO in an atmosphere of zinc vapor and oxygen of which the partial pressures are changed. Adsorbed zinc atoms give rise to donor and oxygen atoms to acceptor surface states. In Fig. 3 are shown p^+ as a function of p_z^+ for $m_z^- = 0$, and m^- as a function of m_z^- for $p_z^+ = 0$. For both curves $\epsilon = 12$,† $T = 600°\text{K}$, $n_\infty^- = 10^{21} m^{-3}$, $p_\infty = 3{\cdot}7 \times 10^5 m^{-3}$, and $\Sigma q^+(\zeta_z^+) = \Sigma q^-(\zeta_z^-) = 0$. This results in the conduction band in the field-free interior being 0·58 V below the Fermi level, that is $\phi_\infty = -0{\cdot}58$ V. As there are several donor levels in ZnO within 0·1 V of the conduction band,[13] the assumption of total donor ionization at 600°K and outside the space-charge layer is realistic. The values for n_∞^-, p_∞, and the corresponding expression for $F(p_z^+/p^+)$ has been derived in the Appendix, where spherical energy surfaces have been assumed. Using this expression and taking for the effective mass the mass of the electron in vacuum, the upper dotted portion of the curve in Fig. 3 was obtained. At the point indicated by $\phi_0 = 0$, the Fermi level at the surface

* This value for the energy gap was obtained by MOLLWO[11] using the absorption edge. A value for the thermal-energy gap seems not to be available.

† This value was obtained by HAHN,[12] using ZnO powder. After the present paper was ready for publication, a value of 8·5 for single crystals was found by T. S. BENNEDICT and A. R. HUTSON of the Bell Laboratories (private communication), using a cavity-perturbation method at a frequency of 24 kMc/s. The effect of this difference on the results presented here is small.

VOL. 4 1958

coincides with the lower edge of the conduction band. The curve for pure oxygen gas is, for low values of m_z^-, the same as that for pure zinc vapor. At higher coverage, this is $\phi_0-\phi_\infty < -4kT/e$, the term $\ln(m_z^-/m^-)$ in equation (10) dominates and the curve shows a nearly horizontal region. The quite different behavior for zinc and oxygen for these values of p_z^+ and m_z^- is peculiar to semiconductors with a large forbidden gap where an inversion layer is not readily established, and is to be contrasted with the behavior of germanium.[14] Thus, in case a depletion layer is formed, for instance with oxygen, the density of the screening charge is limited to the donor density. If additional charge is adsorbed on the surface, overall neutrality can only be maintained by increasing the depth of the space-charge layer. Thus the barrier height is progressively increased, causing the near-horizontal region of the curve for oxygen. For zinc vapor, when Boltzmann statistics may still be applied for the electrons in the conduction band, the density of the screening charge increases rapidly if more electrons are drawn to the surface. Beyond a point where the density of conduction electrons approaches the effective density of states in the conduction band, Fermi statistics has to be applied, and a behavior similar to that for oxygen is observed, as expressed by the logarithmic dependence of p^+ on p_z^+/p^+ in (d) of the Appendix. For a barrier height of about $-32\ kT/e$ the flat region for the oxygen curve ends as the Fermi level nears the top of the valence band and the space-charge region becomes a p-type enrichment layer. After a transition region the contribution of the donors to the screening charge can be neglected, and the remainder of the oxygen curve can be obtained by displacing the upper part of the curve for zinc to the left as follows from the results obtained in the Appendix.

In Fig. 4 is shown the height of the space-charge barrier, $\phi_0-\phi_\infty$, in units of kT/e as a function of the net density of electronic charges at the surface. The lower curve pertains to a positively charged, the upper to a negatively charged surface. The portions of the curve on the right are parallel, with a vertical separation equal to the potential of the top of the valence band with respect to the Fermi level minus the potential of the Fermi level with respect to the conduction band, both in the field-free region of the semiconductor and measured in units

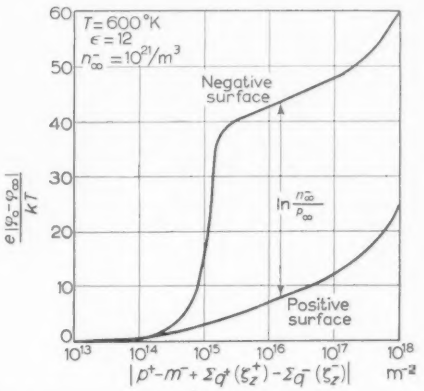


FIG. 4. Dependence of the voltage drop across the space-charge layer in kT/e units on the density of the net surface charge in electronic units for both a positive and a negative surface.

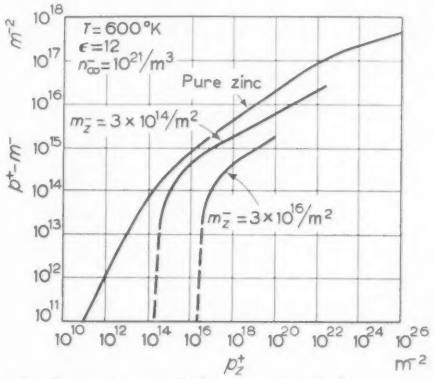


FIG. 5. Dependence of the density of the net surface charge in electronic units on p_z^+, for different values of m_z^- for the range in which the surface is positive.

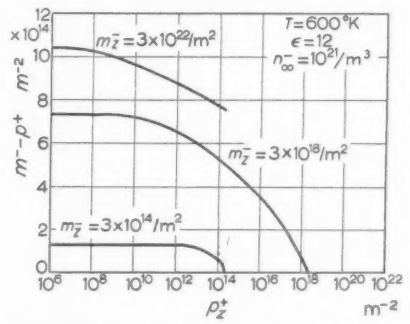


FIG. 6. Dependence of the density of the net surface charge in electronic units on p_z^+ for different values of m_z^- for the range in which the surface is negative.

of kT/e. In Figs. 5 and 6 are shown $p^+ - m^-$ as a function of $p_z{}^+$ for different values of $m_z{}^-$. This pertains to the case in which the semiconductor is surrounded by a mixture of both zinc vapor and oxygen. Use has been made of equation (10) for $\Sigma q^+(\zeta_z{}^+) = \Sigma q^-(\zeta_z{}^-) = 0$. Fig. 5 shows the region for positive, Fig. 6 for negative, net surface charge. The height of the space-charge barrier at any point can be readily obtained with the aid of Fig. 4. The two curves for mixtures in Fig. 5 and the two upper curves in Fig. 6 are terminated at points were the number of ions at the surface equals the number of sites. These points will in fact never be reached, because of the presence of neutral species, as discussed below. As we may expect that the adsorbed zinc and oxygen will react and thus will extend the ZnO lattice, these curves might not be valid up to one monolayer coverage. The figures show that zero surface charge is reached for $p_z{}^+ = m_z{}^-$ or, according to equation (3), if $P_{Zn}/P_{O_2}{}^{\frac{1}{2}} = (K_2/ADK_1)\exp e(\psi_z{}^+ + \psi_z{}^-)/kT$. This is in accordance with the more general observation made in Section 3. The use of a log scale in Fig. 5 and a linear scale in Fig. 6 for the ordinate is a consequence of the different screening mechanisms prevailing for a positive and a negative surface.

It can be seen from these figures that the presence of oxygen reduces the range over which the net charge at the surface will change when the zinc pressure is increased from zero to the value at which monolayer coverage takes place.

The difference of the densities of positive and negative surface ions as a function of gas pressure is readily obtained if equations (2b) and (3) are combined with the results shown in Figs. 3, 5 and 6. In Fig. 7 is shown the density of zinc ions as a function of zinc pressure for pure zinc vapor, with $\Sigma q^+(\zeta_z{}^+) - \Sigma q^-(\zeta_z{}^-) = 0$. For the left solid curve, the electronic level of an adsorbed zinc atom coincides with the bottom of the conduction band. For the curve to the right this level has been taken 0·7 eV lower. The straight sections of these curves at higher coverage follow the dependence derived by HAUFFE and ENGELL.[5] These curves are the same as $p^+(p_z{}^+)$ apart from the factor $DK_1 \exp[-e\psi_z{}^+/kT]\ (1-\sigma/\sigma_0) = p_z{}^+/P_{Zn}$ which is a constant as long as $\sigma \ll \sigma_0$. In order to show that the saturation region, that is $1-(\sigma/\sigma_0) \neq 1$, is coveragewise very narrow, a dotted line has been put in representing p^+ as a function of $DK_1P_{Zn}/1-(\sigma/\sigma_0)$. Thus the shape of the solid lines is determined by the properties of the semiconductor alone, which consists of bulk properties and any distribution of fast states which may happen to be at the surface. On the other hand the properties relating to the interaction of gas and semiconductor determine both the position of the curve along the abscissa and the ion coverage at which Langmuir saturation occurs. By comparing an experimental curve of the change in surface-charge density versus gas pressure with $p^+(p_z{}^+)$, calculated from the known semiconductor bulk properties and the distribution of fast states, the constant $DK_1 \exp[-e\psi_z{}^+/kT]$ can be obtained. The potential $\psi_z{}^+$, as well as D and K_1 can be evaluated in the following way. If $\overline{p^+}$ and $\overline{p}$ are respectively the densities of ions and neutrals at the surface for which adsorption stops, $\overline{p} = \sigma_0 - \overline{p^+}$. Substituting this in equation (2b) yields,

$$D^{-1}\exp\frac{e\psi_z{}^+}{kT} = \frac{\sigma_0 - \overline{p^+}}{\overline{p_z{}^+}} \qquad (11)$$

$\overline{p_z{}^+}$, the value of $\overline{p_z{}^+}$ corresponding with $\overline{p^+}$, is available from the calculated curve $p^+(p_z{}^+)$. Thus $D^{-1}\exp(e\psi_z{}^+/kT)$ is obtained if σ_0 is known and consequently so are D and $\psi_z{}^+$, using results for two values of T. If σ_0 is not known but is independent of temperature, D, $\psi_z{}^+$, and σ_0 can be derived by solving equation (11) for three values of temperature.

In Fig. 8 is plotted the total density of charged and uncharged species as a function of K_1P_{Zn} for pure zinc vapor, $D = 1$, and for the same values of

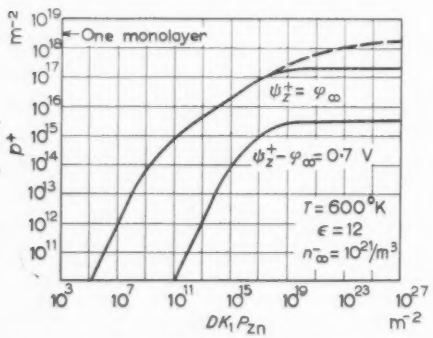


FIG. 7. Dependence of the density of zinc ions at the surface on the pressure of pure zinc vapor for two values of $\psi_z{}^+ - \phi_\infty$.

the energy level used in Fig. 7. When the level is at the bottom of the conduction band, the adsorbed species are practically all ionized up to a coverage

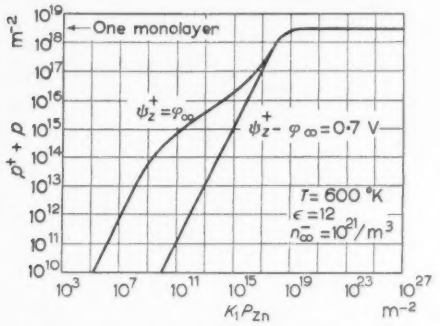


FIG. 8. Dependence of the total density of zinc atoms both ionized and neutral on the pressure of pure zinc vapor for two values of $\psi_z^+ - \phi_\infty$.

of about 10^{-2} monolayer. At saturation, however, only one atom out of seventeen is ionized.

There is some doubt about the validity of the choice $D = 1$. In case $D = \frac{1}{2}$, one monolayer coverage will take place if only one out of about twenty instead of seventeen atoms is ionized. If the level is 0·7 eV below the conduction band and $D = 1$, practically all the adsorbed species are neutral even at the low coverage where space-charge depression does not play a role. At saturation only one out of 1000 adatoms is ionized.

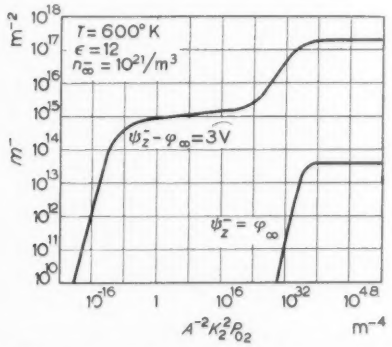


FIG. 9. Dependence of the density of oxygen ions on the pressure of pure oxygen gas for two values of $\psi_z^- - \phi_\infty$.

In Fig. 9 is plotted the density of oxygen ions as a function of $A^{-2}K_2^2P_{O_2}$ for two values of the surface level and $\Sigma\, q^+(\zeta_z^+) - \Sigma\, q^-(\zeta_z^-) = 0$. For the curve on the left the energy level coincides with the top of the valence band. The region for low coverage where the space charge does not play a role, the near horizontal region derived by HAUFFE and ENGELL[5] and corresponding with the sharp increase in the space-charge barrier height and thickness, the p-type inversion region, and the Langmuir saturation region are all shown. Because of the very large range of pressures involved, one can expect to see experimentally only a small part of the curve. For the curve on the right, the energy level has been assumed to coincide with the bottom of the conduction band. The adsorption is stopped before the barrier height of the space-charge layer becomes important for the adsorption.

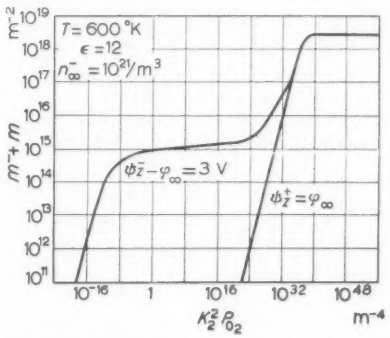


FIG. 10. Dependence of the total density of both oxygen atoms and ions on the pressure of oxygen gas for two values of $\psi_z^- - \phi_\infty$.

In Fig. 10 is plotted the density of neutral atoms and ions combined versus $K_2^2P_{O_2}$ for $A = 1$ and for the same energy values of the surface levels as in Fig. 9. At monolayer coverage, about one out of seventeen adsorbed atoms is ionized if the level coincides with the top of the valence band. The barrier height of the space-charge layer for a certain gas pressure can be readily obtained by combining Fig. 4 with Figs. 7 or 9.

SUMMARY

By assuming a simple picture for the electronic surface levels of adatoms, in which only the change of long-range electrostatic forces with coverage has been taken into account, adsorption isotherms were obtained for pure and mixed gases interacting with a semiconductor. The influence of the constant determining the equilibrium between the gas

phase and the neutral adsorbed species, of the energy of the adatom electronic levels, of fast states, and of the bulk properties of the semiconductor on the shape of the adsorption isotherm for a pure gas, are readily discernible.

APPENDIX

An expression for $F(p_z^+/p^+)$ will be obtained for a high-density electron-enrichment layer at the surface without resorting to the commonly used approximation of a Maxwellian energy distribution of the conduction electrons. The total space charge is so large in this case that the influence of donors can be neglected.

Assuming spherical energy surfaces at the bottom of the conduction band, the concentration of the conduction electrons at a point with potential ϕ is given by

$$n^- = \frac{4\pi}{h^3}(2m^*)^{3/2}\int_{-e\phi}^{\infty}\frac{(E+e\phi)^{\frac{1}{2}}}{1+\exp(E/kT)}\,dE$$

E is the energy of a level in the conduction band relative to the Fermi level. The upper limit of the integral has been chosen as ∞, since a large part of the conduction band is still empty.

Accordingly,

$$F^2(p_z^+/p^+) = \frac{-8\pi\epsilon\epsilon_0}{eh^3}(2m^*)^{3/2}\int_{\phi_0}^{\phi_\infty}d\phi\int_{-e\phi}^{\infty}\frac{(E+e\phi)^{\frac{1}{2}}\,dE}{1+\exp(E/kT)}. \quad (a)$$

Substituting $z = E+e\phi$ and changing the order of integration gives

$$F^2(p_z^+/p^+) = \frac{-8\pi\epsilon\epsilon_0}{eh^3}(2m^*)^{3/2}\int_0^{\infty}z^{\frac{1}{2}}\,dz\int_{\phi_0}^{\phi_\infty}\frac{d\phi}{1+\exp[(z-e\phi)/kT]}$$

$$= \frac{-8\pi\epsilon\epsilon_0 kT(2m^*)^{3/2}}{e^2h^3}\times\left[\int_0^{\infty}z^{\frac{1}{2}}\ln\left(1+\exp\frac{e\phi_\infty-z}{kT}\right)dz-\int_0^{\infty}z^{\frac{1}{2}}\ln\left(1+\exp\frac{e\phi_0-z}{kT}\right)dz\right].$$

We will assume the Fermi level to be at least a few kT below the conduction band in the field-free interior of the semiconductor, so that

$$\ln\left[1+\exp\frac{e\phi_\infty-z}{kT}\right] \approx \exp\frac{e\phi_\infty-z}{kT}$$ and thus

$$F^2(p_z^+/p^+) = \frac{-8\pi\epsilon\epsilon_0 kT(2m^*)^{3/2}}{e^2h^3}\times\left[\frac{\pi^{\frac{1}{2}}(kT)^{3/2}}{2}\exp\frac{e\phi_\infty}{kT}-\int_0^{\infty}z^{\frac{1}{2}}\ln\left(1+\exp\frac{e\phi_0-z}{kT}\right)dz\right]. \quad (b)$$

In order to evaluate the remaining integral, we will consider the two cases $e\phi_0 > 0$ or the Fermi level above the conduction band and $e\phi_0 \leqslant 0$ or the Fermi level at or below the conduction band.

In the first case

$$\int_0^{\infty}z^{\frac{1}{2}}\ln\left[1+\exp\frac{e\phi_0-z}{kT}\right]dz = \int_0^{e\phi_0}\frac{z^{\frac{1}{2}}(e\phi_0-z)}{kT}\,dz+\int_0^{e\phi_0}z^{\frac{1}{2}}\ln\left[1+\exp\frac{-(e\phi_0-z)}{kT}\right]dz+\int_{e\phi_0}^{\infty}z^{\frac{1}{2}}\ln\left[1+\exp\frac{e\phi_0-z}{kT}\right]dz.$$

The first integral can be readily integrated. The logarithm in the integrand of the second and third integral can be expanded in a power series which thus can be integrated term by term. The result substituted in (b) gives for $e\phi_0 > 0$

$$F(p_z^+/p^+) = K^{\frac{1}{2}}\left[-\frac{\pi^{\frac{1}{2}}}{2}\frac{n_\infty^-}{N_c}+\frac{4}{15}\beta_0^{5/2}+\frac{\pi^2}{6}\beta_0^{\frac{1}{2}}+\sum_{n=1}^{\infty}(-1)^n n^{-5/2}\left\{\exp(-n\beta_0)\int_0^{(n\beta_0)^{\frac{1}{2}}}\exp y^2\,dy-\right.\right.$$

$$-\exp(n\beta_0)\int_{(n\beta_0)^{\frac{1}{2}}}^{\infty}\exp -y^2\,dy\Big)\Big]^{\frac{1}{2}}. \qquad \text{(c)}$$

With $\beta_0 = \ln[p_z^+ n_\infty^-/p^+ N_c] = e\phi_0/kT$ according to equation (5), $N_c = 2(2\pi m^* kT/h^2)^{3/2}$ the effective density of states in the conduction band and $K = (8\pi\varepsilon\varepsilon_0/e^2h^3)\times(2m^*)^{3/2}(kT)^{5/2}$. The first term on the right-hand side of (c) can usually be neglected. The expression

$$\exp(-n\beta_0)\int_0^{(n\beta_0)^{\frac{1}{2}}}\exp y^2\,dy$$

as a function of $(n\beta_0)^{\frac{1}{2}}$ has been tabulated by LASH MILLER and GORDON.[15] The second integral in (c) is the complementary error function. The series converge reasonably rapid, so that only a few terms are required. If the Fermi level is a few kT above the conduction band at the surface, (c) can be simplified to

$$F(p_z^+/p^+) = \frac{2}{\sqrt{15}}K^{\frac{1}{2}}\left(\ln\frac{n_\infty^-}{N_c}+\ln\frac{p_z^+}{p^+}\right)^{5/4}. \qquad \text{(d)}$$

For $e\phi_0 \leqslant 0$, the integral in (b) can be integrated term by term after expanding the logarithm. Consequently for $e\phi_0 \leqslant 0$

$$F(p_z^+/p^+) = K^{\frac{1}{2}}\left[-\frac{\pi^{\frac{1}{2}}}{2}\frac{n_\infty^-}{N_c} - \right.$$

$$\left.-\frac{\pi^{\frac{1}{2}}}{2}\sum_{n=1}^{\infty}(-1)^n n^{-5/2}\exp n\beta_0\right]^{\frac{1}{2}}. \qquad \text{(e)}$$

If the Fermi level is at the surface, a few kT below the bottom of the conduction band, (e) can be simplified to $F(p_z^+/p^+) = +\beta^{\frac{1}{2}}(-1+p_z^+/p^+)^{\frac{1}{2}}$, which equals the conduction electron contribution to the screening in equation (9).

For a p-type inversion layer, the space charge is formed by holes. In case the hole density is large enough for the influence of the donors to be neglected, the expression for $F(p_z^+/p^+) = F(m^-/m_z^-)$ differs from (c) and (e) only in that m^* is the effective mass of a hole, that n_∞^-/N_c is replaced by p_∞/N_v and β_0 by

$$\ln\left[\frac{m_z^-}{m^-}\cdot\frac{p_\infty}{N_v}\right]$$

as can be easily seen by comparing (a) with the corresponding expression for the hole-density if the Fermi level is close to or in the valence band at the surface. N_v is the effective density of states in the valence band.

NOMENCLATURE

P_d	partial pressure donor gas
r	number of atoms per molecule of donor gas
P_a	partial pressure acceptor gas
s	number of atoms per molecule of acceptor gas
n_0^-	concentration of conduction electrons at the surface
K_1	equilibrium constant donor gas $\rightleftarrows$ donor adatoms
σ_0	total number of adsorption sites
σ	number of adsorption sites occupied by both atoms and ions
$[(1-(\sigma/\sigma_0)]_d$	fraction of donor adsorption sites which is vacant
K_2	equilibrium constant acceptor gas $\rightleftarrows$ acceptor adatoms
$[(1-(\sigma/\sigma_0)]_a$	fraction of acceptor adsorption sites which is vacant
p	density donor adatoms
p^+	density donor adions
p_z^+	density donor adions if the barrier height of the space-charge layer is neglected
D	degeneracy electronic level donor adion
ψ^+	potential electronic level donor adion
ψ_z^+	potential electronic level donor adions if barrier height of the space-charge layer is neglected
m	density acceptor adatoms
m^-	density acceptor adions
m_z^-	density acceptor adions if barrier height of the space-charge layer is neglected
A	degeneracy electronic level acceptor adion
ψ^-	potential electronic level acceptor adion
ψ_z^-	potential electronic level acceptor adion if barrier height of the space-charge layer is neglected
e	absolute value of electronic charge
k	Boltzmann constant
T	absolute temperature
ϕ	potential of the lower edge of the conduction band with respect to the Fermi level
ϕ_0	value of ϕ at the surface
ϕ_∞	value of ϕ in the field-free interior
$q(\zeta_z^+)$	total density of charged and neutral donor-type fast states
ζ^+	potential electronic level donor-type fast state
ζ_z^+	potential electronic level donor-type fast state if the barrier height of the space-charge layer is neglected
$D(\zeta_z^+)$	degeneracy electronic level donor-type fast state
$q^+(\zeta_z^+)$	density of charged donor-type fast states
$q(\zeta_z^-)$	total density of charged and neutral acceptor-type fast states
ζ_z^-	potential electronic level acceptor-type fast states if the barrier height of the space-charge layer is neglected
$A(\zeta_z^-)$	degeneracy electronic level acceptor-type fast state

$q^-(\zeta_z^-)$ density of charged acceptor-type fast states
x distance of a point in the crystal from the surface
ε dielectric constant
ε_0 $= 8{\cdot}85\times10^{-12}$ F/m permittivity of vacuum
$\rho^+(\phi)$ charge density caused by positive impurities in the space-charge layer
$\rho^-(\phi)$ charge density caused by negative impurities in the space-charge layer
n^- density of conduction electrons
n_∞^- density of conduction electrons in the field-free interior of the semiconductor
p_∞ density of holes in the field-free interior
N_{di} donor density
$1/\alpha_i$ ionized fraction of donors in field-free interior
N_{ai} acceptor density
$1/\beta_i$ ionized fraction of acceptors in field-free interior
$\beta = 2kT\varepsilon\varepsilon_0 n_\infty^-/e^2$
P_{Zn} pressure zinc vapor
P_{O2} pressure oxygen vapor
$\bar{p}$ density of donor adatoms at one monolayer coverage
$\overline{p^+}$ density of donor adions at one monolayer coverage
$\overline{p_z^+}$ value of p_z^+ corresponding with $\overline{p^+}$
E energy of an electronic state in the conduction band with respect to the Fermi level
h Planck's constant
m^* effective mass of an electron in the conduction band
$z = E+e\phi$
$\beta_0 = e\phi_0/kT$
N_C effective density of states in the conduction band
$K = 8\pi\varepsilon\varepsilon_0/e^2h^3(2m^*)^{3/2}(kT)^{5/2}$
N_v effective density of states in the valence band

REFERENCES

1. Brattain W. H. and Bardeen J. *Bell Syst. Tech. J.* **32,** 1 (1953).
2. Morrison S. R. *J. Phys. Chem.* **57,** 860 (1953).
3. Bardeen J. *Phys. Rev.* **71,** 717 (1947).
   Weiss P. B. *J. Chem. Phys.* **21,** 1531 (1953).
4. Aigrin P. and Dugas C. *Z. Electrochem.* **56,** 363 (1952).
5. Hauffe K. and Engell H. J. *Z. Electrochem.* **56,** 366 (1952).
   Engell H. J. and Hauffe K. *Z. Electrochem.* **57,** 762 (1953).
   Hauffe K. *Angew. Chem.* **67,** 189 (1955).
6. Thomas D. G. and Lander J. J. *J. Phys. Chem. Solids* **2,** 318 (1957).
7. Garrett C. G. B. and Brattain W. H. *Phys. Rev.* **99,** 376 (1955).
   Stats H., de Mars G. A., Davis L. and Adams A. *Phys. Rev.* **101,** 1272 (1956).
8. Schrieffer J. R. *Phys. Rev.* **97,** 641 (1955).
9. de Boer J. H. *The Dynamical Character of Adsorption.* Clarendon Press, Oxford (1953).
10. Schottky W. and Spenke E. *Wiss. Veröff. Siemens-Werken* **18,** 225 (1939).
    Shockley W. *Bell Syst. Tech. J.* **28,** 435 (1949).
    Kingston R. H. and Neustadter S. F. *J. Appl. Phys.* **26,** 718 (1955).
11. Mollwo E. *Z. Angew Phys.* **6,** 257 (1954).
12. Hahn E. E. *J. Appl. Phys.* **22,** 855 (1951).
13. Hutson A. R. *Bull Amer. Phys. Soc.* **1,** 381 (1956).
14. Brown W. L. *Phys. Rev.* **91,** 518 (1953).
15. Lash Miller W. and Gordon A. R. *J. Phys. Chem.* **35,** 2878 (1931).

VOL 4 195

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 91–100.

THE INTERACTION OF OXYGEN WITH CLEAN SILICON SURFACES

J. T. LAW

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

(*Received* 27 *June* 1957)

Abstract—An experimental method for measuring the kinetics of the take-up of oxygen by an evaporated silicon film is described. The nature of the evaporated film as a function of its thickness is discussed. The effect of temperature and oxygen pressure on the kinetics of the oxidation process were investigated, and it was found that the rate was given by $dn/dt = ap^{\frac{1}{2}} \exp(-bn)$, where a and b are constants, p is the oxygen pressure, and n is the number of oxygen molecules taken up at time t. All the data reported are for oxygen pressures in the range $10^{-4}-4\times10^{-2}$ mm Hg.

VOL. 4 1958

INTRODUCTION

BECAUSE of the interest in the surface electrical properties of silicon and the effects of adsorbed gases, a knowledge of the surface oxidation process at room temperature is of considerable importance. Previous measurements on this system consist of two types. GREEN and KAFALAS[1] crushed a sample of silicon in vacuum and followed the subsequent uptake of oxygen by means of the pressure change in a constant-volume system. They report (at least germanium and we may assume the same to be true for silicon, as the crystals are grown under comparable conditions) that considerable quantities of gas are released during the crushing process, but they argue that the pressure rise is so small that the degree of surface contamination must be less than 10^{12} molecules cm^{-2}. The oxidation kinetics found were described by the equation

$$N = (0{\cdot}072)(11{\cdot}85\times10^{14})\log t + 11{\cdot}85\times10^{14} \quad (1)$$

where N is the number of oxygen atoms per cm^2 removed from the gas phase in a time t (min). This is essentially an Elovich-type equation without the additional constant in the log t term.

The other type of measurement, made by ARCHER,[2] involved the use of elliptically polarized light to determine the kinetics of the oxidation process at atmospheric pressure on an etched silicon surface. Here also the kinetics were described by an Elovich-type equation for film thicknesses greater than one layer of oxygen atoms. We are interested in investigating both regions of the oxidation process: (i) the fast chemisorption at very low oxygen pressures and (ii) the subsequent growth of an oxide film. Only the latter has been studied by previous workers.

Ideally this type of measurement should be made on an atomically clean face of a single crystal. There are available two convenient methods for producing a clean silicon surface. The first consists of heating a sample to temperatures above 1550°K[3] in high vacuum, while the second is one of positive-ion bombardment. In either of these a considerable amount of silicon is evaporated or sputtered onto the walls of the tube, so that the apparent area of this material far exceeds that of the original filament. Nevertheless, by carrying out such an evaporation process in high vacuum one can obtain a high-area, clean, silicon surface whose principal disadvantage is that it is of unknown crystalline orientation. All the measurements reported in the present work were made on such a high-vacuum-evaporated silicon surface.

In this paper we will describe the oxidation kinetics for pressures between 10^{-4} and 4×10^{-2} mm Hg at temperatures of 273, 298 and 323°K. The nature of the evaporated films as a function of thickness will also be discussed.

EXPERIMENTAL

The oxidation kinetics were measured by following the rate of pressure change in a constant-volume

system. As we will see later, the rate depends on both the film thickness and the oxygen pressure, so that corrections have to be made for the pressure change during the course of a run. The apparatus consisted essentially of a tube containing a silicon filament, a Pirani gauge, and a gas-dosing system consisting of vacuum valves.[4] The whole system could be baked out under vacuum, and pressures of 5×10^{-10} were readily attained. Spectroscopically pure oxygen was used without further purification.

The silicon filament was degassed by a series of high-temperature flashes. The bulb surrounding it was then cooled to 273°K and evaporation carried out while a high vacuum was maintained. The filament temperature during evaporation was 1610 ± 10°K, as determined by an optical pyrometer, and included an emissivity correction of 70°K.[5] During this stage the pressure never rose above 10^{-8} mm Hg. The normal time of evaporation was from 30 to 400 sec with the filament at 1610°K. By varying the evaporation time, films of different thicknesses could be produced, the number of silicon molecules in each being simply calculated from the known vapor-pressure data of silicon.[6] After evaporation the film was pretreated at 50°C for 300 sec to render it stable during subsequent chemisorption. A considerable quantity of heat is liberated in the initial stage of the chemisorption process, and it is important that no further sintering occurs during this stage. We have found that the area (which is a measure of the degree of sintering) of the silicon film is essentially unchanged in going from 273 to 323°K. There is an apparent slight decrease, but it amounts to less than 5 per cent of the total area.

After sintering, the film was cooled to the required temperature, a measured dose of gas admitted, and the pressure in the system followed by the Pirani gauge. By using a large number of small doses, it was possible to study the effect on the oxidation rate of both time and pressure. The lowest pressure conveniently measured in the system was 10^{-5} mm Hg, so that no details of the fast chemisorption could be obtained.

Previous work[7] has shown that the "fast" oxygen uptake at room temperature can be used to measure the area of an evaporated film. By "fast" uptake is meant that which occurs at pressures of the order of 10^{-5} mm Hg and in times of less than 1–2 min. By using this method one can utilize the initial stages of the oxidation process to characterize the film area by assuming that the chemisorbed layer consists of one oxygen atom for each surface silicon atom.

RESULTS

(i) *Variation of surface area with amount evaporated*

No absolute method of determining the areas of films to within 20 per cent is available. A BET area determination would have been valuable, but the large dead space of the present apparatus makes it unsuitable for such measurements. We have, therefore, followed previous workers in comparing areas by using the amount of "fast" chemisorption.

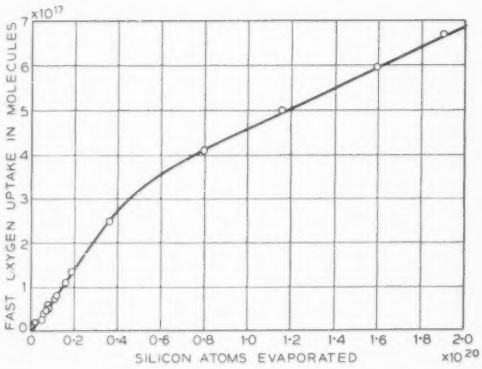


FIG. 1. The variation of fast oxygen uptake (which is proportional to the film area) as a function of the number of silicon atoms evaporated to form the film.

Fig. 1 is a plot of this uptake (in molecules) against the number of silicon atoms evaporated for a series of measurements. It can be seen that above a certain thickness a linear relationship exists between the area, which is proportional to the number of oxygen molecules in the "fast" uptake region, and the amount evaporated. Similar results have been reported for iron films.[8] On the other hand BEECK *et al.*[9] found that for palladium films a linear plot passing through the origin was obtained. For thin films, we find a considerable divergence from the line that characterizes relatively thick films. Any discrepancy between our data and those of BEECK *et al.* could be explained by the fact that we have investigated films one-tenth the thickness of his, and it is only in this region that any difference is to be noted. One

could also ascribe the apparently low areas of the thicker films to sintering. This could be caused by either (i) the heat of chemisorption liberated when a large quantity of oxygen is adsorbed or (ii) the radiant heating of the film by the filament. Obviously for thicker films the evaporating time is longer, so that the wall temperature is higher. We have eliminated these possibilities by evaporating the film in a series of short-time flashes and also by admitting the oxygen in small doses. By admitting the oxygen in small doses, we are essentially using the same conditions as obtained for the thin-film region of Fig. 1. Even if the thermal conductivity of the film is low, a smaller amount of sintering should occur for the case of slow oxygen admission. The results are given in Table 1.

Table 1. *The effect on the film area of varying the evaporating conditions and rate of oxygen admission*

Condition	Area
One 120-sec flash+fast O_2 admission	$5{\cdot}3\times10^{17}$
Three 40-sec flashes+fast O_2 admission	$5{\cdot}2\times10^{17}$
Six 20-sec flashes+fast O_2 admission	$5{\cdot}1\times10^{17}$
One 120-sec flash+slow O_2 admission	$5{\cdot}1\times10^{17}$

(ii) *The oxidation of silicon at constant temperature*

When oxygen was admitted to a silicon film which had been presintered at 323°K and later equilibrated at 273, 298, or 323°K, about 50 per cent of the quantity of gas finally taken up was chemisorbed within 1–2 min and with a residual pressure of about 10^{-5} mm Hg. After this fast uptake, a slow removal of oxygen from the gas phase occurred, which continued for at least 24 hr at a measurable rate.

Subsequent to the fast uptake, doses of oxygen were chosen so that the pressure in the system was higher than 5×10^{-4} mm 1 min after admission. All rate measurements reported in the present paper, therefore, refer to films which were covered by a chemisorbed monolayer before the first measurement was made. Preliminary measurements showed that the rate was controlled by the amount already adsorbed and also by the pressure, so that the following method was used to investigate the pressure-dependence.

A small dose of gas was admitted and a pressure time plot recorded in this low-pressure (P_1) region for 10–30 min; a further dose was then introduced at a higher pressure (P_2) and the rate remeasured. Both curves in the P_1 and P_2 regions could then be extrapolated to the time of admission of the second dose, and the two rates R_1 and R_2 at the two pressures P_1 and P_2 obtained. The values R_1 and R_2 then refer to constant surface coverage but different oxygen pressures. By comparing the pressures P_1 and P_2 with the volumes of the various parts of the system, it was readily shown that no instantaneous adsorption occurred, so that R_1 and R_2 do indeed correspond to a fixed value of surface coverage. Values of the exponent n calculated from

$$R_2/R_1 = (P_2/P_1)^n \qquad (2)$$

are given in Table 2. From the listed values of n, one concludes that the rate of oxidation varies as the square root of the oxygen pressure for constant coverage and temperature. This implies that dissociation of the adsorbed oxygen molecule occurs in the rate-determining reaction.

A typical plot of the variation of pressure in the

Table 2. *Film presintered at* 323°K, *oxygen adsorption at* 298°K. *Rates in molecules/sec*

Pressure		Rate		
P_1	P_2	R_1	R_2	n
$3{\cdot}87\times10^{-4}$	$2{\cdot}08\times10^{-3}$	$1{\cdot}08\times10^{13}$	$2{\cdot}76\times10^{13}$	0·56
$3{\cdot}18\times10^{-3}$	$6{\cdot}60\times10^{-3}$	$3{\cdot}36\times10^{13}$	$4{\cdot}68\times10^{13}$	0·45
$5{\cdot}25\times10^{-4}$	$3{\cdot}61\times10^{-3}$	$2{\cdot}76\times10^{13}$	$7{\cdot}20\times10^{13}$	0·53
$1{\cdot}81\times10^{-3}$	$3{\cdot}53\times10^{-3}$	$3{\cdot}00\times10^{12}$	$4{\cdot}30\times10^{12}$	0·53
			Average:	0·52

system with time at 298°K for several gas doses is shown in Fig. 2. Both the rate of adsorption and the coverage were calculated from the pressure readings, so that the type of curve shown reflects the change in adsorption rate due to both increasing coverage and decreasing pressure. One can readily eliminate the latter, as we have already shown that the rate is directly proportional to the square root of the pressure. The type of curve shown in Fig. 2 is described well by the equation

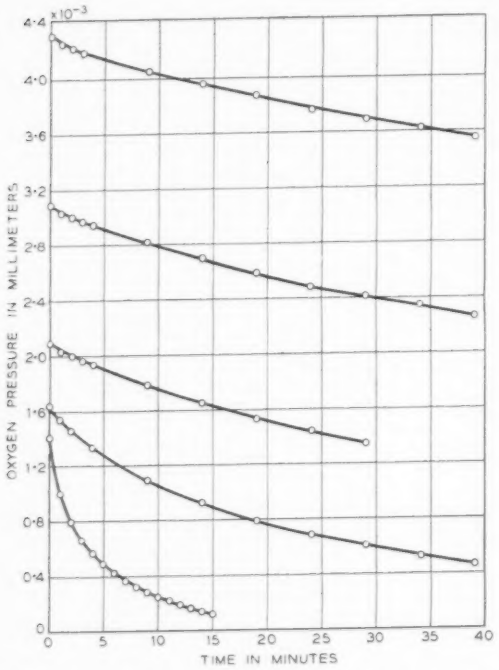


FIG. 2. Some typical curves showing the change in pressure in the system as a function of time for a number of doses. The decrease in rate with increasing time reflects both the increased surface coverage and the lowered oxygen pressure over the film.

$$\frac{dn}{dt} = ap^{\frac{1}{2}} \exp(-bn) \qquad (3)$$

where n is the number of atoms reacted in time t, p is the oxygen pressure, and a and b are constants. Hence

$$\ln[(dn/dt)/p^{\frac{1}{2}}] = \ln a - bn \qquad (4)$$

so that the quantity on the left-hand side is a linear function of n. In Fig. 3 data from several runs at the three temperatures are plotted in this way, resulting in linear curves.

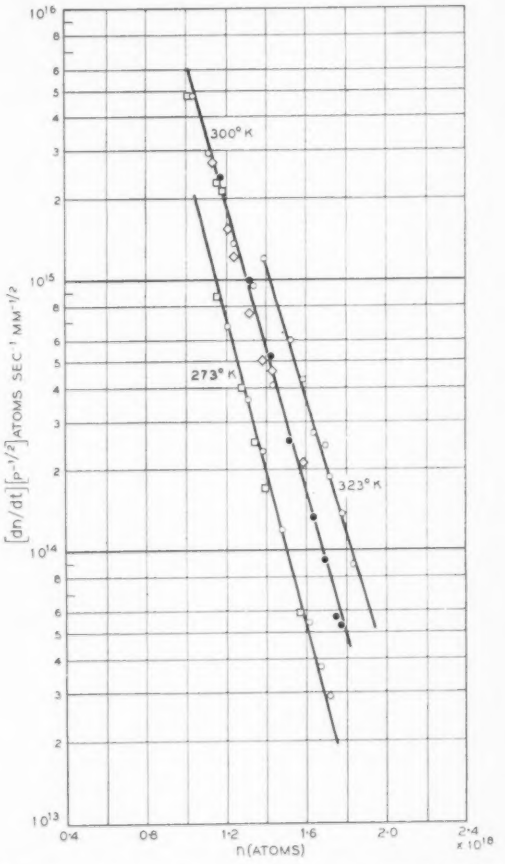


FIG. 3. The rate of oxygen removal plotted against coverage according to equation (3). Data are shown for several runs at the three temperatures 273, 300 and 323°K.

(iii) *The effect of temperature at constant coverage and pressure*

We have investigated the effect of temperature in two ways. The first was to evaporate a series of films under constant conditions of time and filament temperature, so as to obtain essentially identical films, and then to carry out measurements as described above to obtain dn/dt as a function of pressure and coverage. This gave us the results

already shown in Fig. 3. The temperature coefficient was then checked by evaporating a single film, admitting oxygen to it at 273°K, and following the rate for about 30 min, after which the temperature was raised to 298°K and subsequently to 323°K, at each of which temperatures a new rate was determined. The results of a typical run of this type are

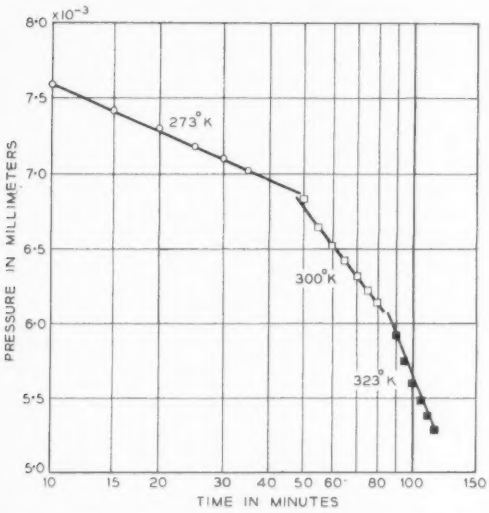


FIG. 4. The effect of changing temperature during a single run, showing the increased rate at elevated temperatures.

shown in Fig. 4. From the ratio of the slopes of the curves immediately before and after a change in temperature, the effect of temperature on the rate at constant pressure and film thickness can be calculated. If R_{T_1} and R_{T_2} are the rates at temperatures T_1 and T_2, then the activation energy is given by

$$-\ln\left[\frac{R_{T1}}{R_{T2}}\right] = \frac{\Delta E^*}{R}\left(\frac{1}{T_1}-\frac{1}{T_2}\right) \qquad (5)$$

At approximately one monolayer coverage, this gives $\Delta E^* = 5{\cdot}8$ kcal/mole, which may be compared with the value of 5·5 kcal/mole obtained by GREEN *et al.*[10] for the germanium–oxygen system. Over the range of coverage investigated ($\theta = 1 \rightarrow 2$), the activation energy increases linearly up to approximately 9·4 kcal/mole.

(iv) *The effect of light on the rate of oxidation*

During runs at 298°K the sample was alternately placed in a dark container and then exposed to intense illumination. No detectable change in the rate of take-up was observed.

(v) *The effect of evacuation on the rate of oxidation*

Some experiments were carried out to determine the effect of stopping the oxidation run, pumping out the system, and then readmitting oxygen. The first of these consisted of adsorbing oxygen to a coverage of $\theta = 1{\cdot}5$ (where $\theta = 1$ corresponds to the fast uptake), evacuating the system to 5×10^{-10} mm for 16 hr, and then readmitting oxygen. The new rates were found to be almost twice as great as those to be expected if no change had occurred on the surface while it remained in a vacuum. This presumably meant that either (i) desorption or (ii) some molecular rearrangement had taken place. The first of the two possibilities was eliminated by measuring the pressure increase in the closed system after the cessation of a run followed by evacuation. After suddenly lowering the pressure to 10^{-7} mm, no pressure increase greater than 2×10^{-8} mm occurred over a 30 min period; whereas if the increased oxidation rate had been due to desorption, a pressure increase of at least 7×10^{-5} mm would have been expected in this time interval.

DISCUSSION

(i) *The nature of the silicon film*

In Fig. 1 the area of the film is shown as a function of the amount of silicon evaporated. Similar data have previously been reported for iron,[8] nickel[9] and palladium[11] films. The results obtained on nickel and palladium could be described by a single straight line passing through the origin, so that it has been stated that the area of a film is a linear function of its weight. This implies that throughout the film thickness investigated, the particles of which the films are composed are of constant size and change only in number. Electron-diffraction studies,[12] however, on silver films have shown that as the thickness increases from 30 to 550 Å, the particle size increases from 100 to 5000 Å. If this were indeed the case, one would not expect to obtain a single straight line passing through the origin. Our results are similar to those obtained on an iron film by PORTER and TOMPKINS.[8] These workers, investigating films considerably thicker than ours, obtained a linear

relation between area and volume evaporated which did not pass through the origin. This presumably corresponds to the second linear region of our curve. From our data and also from previously published work, it is possible to estimate the number and size of the particles that make up the film if one assumes a certain constant shape for each particle.

Presumably the two most reasonable shapes from surface-energy considerations are spheres and oblate spheroids. A considerable amount of optical theory of metal films is based on the assumption of spheres, but SCHOFFER[13] has claimed that an ellipsoidal shape is more nearly correct. We will attempt to calculate the size and number of particles for these two cases, assuming uniformity of particle size.

In the case of spheres the area/volume ratio is given by $3/r$, where r is the radius. Hence, we can readily calculate r and therefore also the number of such particles required to make up the measured volume (or area). The results for various elements are listed in Table 3, and in the case of silicon they are plotted in Fig. 5.

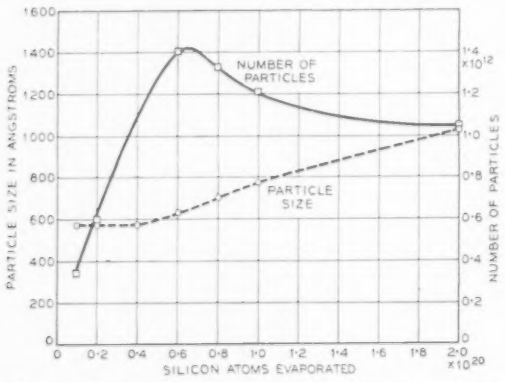


FIG. 5. The number and size of the particles in the evaporated film as a function of film thickness.

Table 3. Particle sizes for various evaporated films

Element	Film weight (mg)	Particle size (Å)
Iron	10–70	310–420
Nickel	10–40	240
Palladium	20–60	710
Silicon	0·2–10	575–1000

If one assumes a particle shape of an oblate spheroid whose major axis is twice the minor axis, then the surface area reduces to $3a^2$, where a is equal to half the major axis, or a surface-to-volume ratio of $2/a$. This simply means that under these conditions a is one and a half times as large as the values listed in Table 3 for r, so that the curves shown in Fig. 5 are essentially unchanged in shape.

If one calculates the number of spheres and oblate spheroids that could be accommodated in a single layer on the available glass wall area, one obtains $1{\cdot}8\times10^{12}$ and $0{\cdot}8\times10^{12}$ respectively for a square array of particles. This assumes a sphere diameter of 1150 Å and a spheroid major axis of 1724 Å. For these small particles we have measured a maximum density of $1{\cdot}4\times10^{12}$, so that it would appear that spheres of silicon of a fixed size form on the glass walls until their density is such that two spheres touch, these then coalesce (with a consequent reduction in the free energy of the system) and, as the curves in Fig. 5 show, the number of particles in the system then decreases slightly while the average size of the particles increases.

(ii) *The area of the film*

As described in the results section, we were forced by the geometry of the system to use the fast uptake to characterize the area of the film. The fast uptake was measured as a number of oxygen atoms, and we need to know (i) whether the oxygen is adsorbed as molecules or atoms and (ii) the number of available sites per cm² of surface. The first of these is readily decided, since GREEN and KAFALAS[1] found a fast uptake on silicon of $1{\cdot}18\times10^{15}$ atoms/cm². This can mean only that the oxygen adsorption is occurring as roughly one atom per surface silicon atom. We can check this by calculating the number of surface silicon atoms for each of the three principal planes. The results obtained, assuming a lattice constant of 5·43 Å,[14] are:

Planes	Sites/cm²
(111)	$7{\cdot}8_5\times10^{14}$
(100)	$6{\cdot}8\times10^{14}$
(110)	$9{\cdot}7\times10^{14}$

While the distribution of crystal faces on an evaporated film is not known, we can say that the number of sites is probably about 8×10^{14}/cm². In

view of the discrepancy between this and the value of $11{\cdot}8\times10^{14}$ obtained by GREEN,[1] we have taken a round figure of 10^{15} sites/cm^2, which should not be in error by more than 20 per cent, i.e. it is of approximately the same accuracy as the BET. This means that our film areas were in the range 20–1600 cm^2; and essentially all the rate data reported were on a film of 10^3 cm^2 area.

(iii) *The logarithmic time-dependence*

The rate of uptake of oxygen (above $\theta = 1$) by evaporated silicon between 273 and 323°K decreases exponentially with increasing amount of gas adsorbed, as described by equation (3)

$$dn/dt = ap^{\frac{1}{2}}\exp(-bn).$$

This rate law has been found to hold very widely in both chemisorption and oxidation. A review of a large amount of the experimental data fitted by this equation has been given by LANDSBERG,[15] who also derived it on the assumption that the surface area which a particle invalidates on adsorption is different from its impact area. The simplest basis for equation (3) is that the activation energy for the process increases linearly with coverage, i.e.

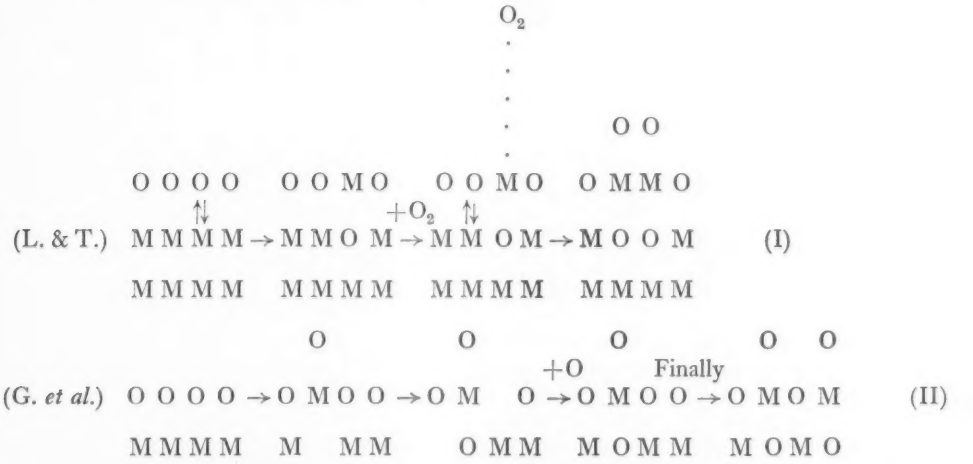
$$\Delta E^* = E_1+nE_2 \qquad (6)$$

PORTER and TOMPKINS[8] explained their results on the hydrogen-iron system in essentially this way. They suggested that the slow step represented the migration of chemisorbed hydrogen atoms from active to inactive surface sites, thereby freeing the active sites for further rapid chemisorption. The activation energy of migration was assumed to increase linearly with coverage. LANYON and TRAPNELL[16] studied a series of metal–oxygen systems and claimed that the migration mechanism was unlikely because the rate of removal of a particular dose was not increased by leaving the oxygenated surface in vacuum for several hours before admitting the dose. This, however, is just what we have found. If an oxygenated silicon surface is left in a vacuum, no gas is desorbed (as determined by pressure measurements), but the rate of take-up of the next dose is considerably increased. This fits in with the picture of some sort of rate-determining molecular rearrangement occurring on the surface. This migration can apparently occur in either the presence or absence of oxygen and leads to the exposure of more active surface sites.

Let us now consider how the migration of adsorbed atoms can expose sites for further adsorption. The two most likely mechanisms are:

(1) If the activation energy for adsorption on a given site is greater than the activation energy for mobility, then those sites will be filled predominately by adsorbed atoms moving from other sites which have low activation energies for adsorption.

(2) Mechanisms involving penetration of the lattice have been advanced by LANYON and TRAPNELL[16] for metals and by GREEN *et al.*[10] for germanium. The postulated mechanisms are shown below:


O2
O O
O O O O O O M O O O M O O M M O
+O2
(L. & T.) M M M M → M M O M → M M O M → M O O M (I)
M M M M M M M M M M M M M M M M
O O O
+O
Finally
(G. et al.) O O O O → O M O O → O M O → O M O O → O M O M (II)
M M M M M M M O M M M O M M M O M O

Both involve the movement of a substrate atom relative to its neighbors. In (I) it exchanges position with an oxygen atom, while in (II) a whole MO unit moves, the cation vacancy being filled with an oxygen atom (or ion). If the driving force for (II), as suggested by GREEN, is the strong repulsion between adjacent $\begin{bmatrix}\mathrm{OO}\\ \mathrm{MM}\end{bmatrix}$ groups as compared with $\begin{bmatrix}\mathrm{OM}\\ \mathrm{MO}\end{bmatrix}$ groups, then there is no *a priori* reason why this process should not proceed on a partly oxygenated surface in vacuum as required by our experimental results. Corroborative evidence for a mechanism such as (II) is found in the work-function data of DILLON[17] and FARNSWORTH.[18] They found no change in the work function of a (110) plane of germanium in going from oxygen pressures of 10^{-5} to 10^{-4} mm and higher, and it does not seem unreasonable that the work function of the initial and final states shown above should be nearly identical. The first mechanism discussed (migration) should lead to an increased work function. From the increased take-up rate after 16 hr. in vacuum, we find that the velocity of rearrangement in vacuum is about one-tenth as fast as that occurring in the presence of oxygen. The oxygen data showed that the rate varied as the square root of the pressure, and this is consistent with the view that oxygen is present on the surface as atoms. A similar dependence on pressure has been found[16] for the interaction of oxygen with molybdenum, rhodium, tungsten, tantalum and zinc.

The present data are not fitted by the CABRERA-MOTT[19] theory for the formation of very thin oxide films. By assuming that the rate-determining step in this region was the removal of an ion from the metal lattice, they derived the equation

$$dn/dt = a\exp(b/n).$$

Obviously, data obeying the Elovich equation will only be fitted by this when n is very small. This may be compared with the results obtained for the oxygen–iron and oxygen–copper systems,[16] where even in the $\theta = 1$–2 range the kinetics obeyed the CABRERA-MOTT equation.

Our picture of the silicon-oxygen system, therefore, is one of fast chemisorption up to a monolayer of atoms, followed by a slow logarithmic process whose rate-determining step includes an atomic-exchange mechanism.

Before we can integrate equation (3)

$$dn/dt = ap^{\frac{1}{2}}\exp(-bn),$$

we need to know the boundary conditions. If the initial process of chemisorption is not described by this equation, then the required condition is that $n = n_m$ at $t = 0$. This assumes that the adsorption process is very fast. Integrating the above equation with this boundary condition, we get

$$n = \frac{1}{b}\ln(ap^{\frac{1}{2}}bt + c) \qquad (7)$$

where $c = \exp(bn_m)$. For the data shown in Fig. 3, a monolayer of atoms corresponds approximately to a total uptake of 10^{18} atoms, and we need to demonstrate that the rate of uptake below this coverage is much faster than the logarithmic equation would predict. For one run at 300°K the coverage was increased from $\theta = 0{\cdot}6$ to $\theta = 0{\cdot}9$ with a value of $(dn/dt)p^{-\frac{1}{2}}$ of approximately 10^{18} atoms $\mathrm{sec}^{-1}\ \mathrm{mm}^{-\frac{1}{2}}$ compared to the value 2×10^{16} expected from the linear region of the curve shown in Fig. 3. Hence equation (7) appears to be applicable.

An alternative way of writing (7), which may make the existence of two separate processes more obvious, is

$$\theta = \frac{1}{b'An_m}\ln[a'b'tp^{\frac{1}{2}} + \exp(b'An_m)] \qquad (8)$$

$$\theta = 1 + \frac{1}{b'An_m}\cdot\ln\left[1 + \frac{a'b'tp^{\frac{1}{2}}}{\exp(b'An_m)}\right]. \qquad (9)$$

These equations, which are derived in the Appendix, give the variation in surface coverage (θ) with time. Here the constants a' and b' refer to total oxygen uptake on the film (rather than to atoms/cm^2), A is the area of the film, and n_m is the number of atoms of oxygen in a monolayer. Thus the product An_m is equal to the total number of oxygen atoms taken up at $\theta = 1$.

From the slopes of the curves in Fig. 3, we can calculate b' and from the intercepts on the ordinate at $\theta = 0$, values of a'. These are listed below in Table 4, together with values of c' calculated from $c' = \exp(b'An_m)$, assuming An_m equal to 8×10^{17} and 1×10^{18}.

Therefore, the measured uptakes (θ) at time t (in

Table 4. Values of the constants a′, b′, *and* c′

Temp.(°K)	a'	b'	c'	
			$An_m = 8\times10^{17}$	$An_m = 1\times10^{18}$
273	$2{\cdot}4\times10^{18}$	$6{\cdot}7\times10^{-18}$	213	812
300	$2{\cdot}5\times10^{18}$	$6{\cdot}1\times10^{-18}$	132	445
323	$2{\cdot}5\times10^{18}$	$5{\cdot}7\times10^{-18}$	96	300

sec) are given for the three temperatures by the following equations:

(a) Assuming $An_m = 8\times10^{17}$

$$\theta_{273} = 0{\cdot}187\ln(16{\cdot}1p^{\frac{1}{2}}t+213) \quad (10)$$

$$\theta_{300} = 0{\cdot}205\ln(15{\cdot}3p^{\frac{1}{2}}t+132) \quad (11)$$

$$\theta_{323} = 0{\cdot}220\ln(14{\cdot}2p^{\frac{1}{2}}t+96) \quad (12)$$

(b) Assuming $An_m = 10^{18}$

$$\theta_{273} = 0{\cdot}149\ln(16{\cdot}1p^{\frac{1}{2}}t+812) \quad (10a)$$

$$\theta_{300} = 0{\cdot}164\ln(15{\cdot}3p^{\frac{1}{2}}t+445) \quad (11a)$$

$$\theta_{323} = 0{\cdot}175\ln(14{\cdot}2p^{\frac{1}{2}}t+300) \quad (12a)$$

Equations are given for the two values of An_m, as there is some uncertainty as to the exact point of completion of the monolayer. The true value of An_m must, however, lie between 8×10^{17} and 10^{18}, being probably much closer to the latter figure.

The data calculated from the above equations at 300°K and at 10^{-2} mm are plotted in Fig. 6, together with values calculated from GREEN's[1] data.

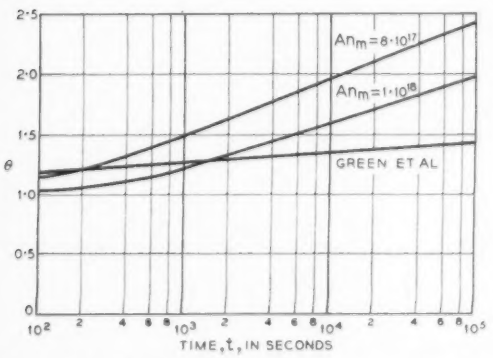


FIG. 6. Rates of uptake of oxygen by silicon at 300°K as a function of time. The upper two curves show the limits of the present data, while the lower curve was calculated from the equation of GREEN *et al.*

Unfortunately we do not know at exactly what pressure his measurements were made, but it must have been in the range $10^{-4}-10^{-2}$ mm Hg.

It can be seen that our rates are considerably higher than his (approximately by a factor of seven). This may be because our surface was initially cleaner. It is difficult to ascribe the difference between the two sets of data to some property of the evaporated film such as edge effects on the small particles of which the film is composed as GREEN's data were also taken on a mass of small particles.

(iv) *The effect of temperature*

As we have seen above, the rate of removal of oxygen from the gas phase is given by equation (3)

$$\frac{dn}{dt} = ap^{\frac{1}{2}}\exp(-bn).$$

Now a condition for the existence of this equation is that the activation energy increase linearly with coverage, i.e.

$$\Delta E^* = E_0+nE_1 \quad (13)$$

or

$$\frac{dn}{dt} = a''p^{\frac{1}{2}}e^{-E_0/RT}e^{-nE_1/RT} \quad (14)$$

$$\ln\left(\frac{dn}{dt}\right) = \ln a''p^{\frac{1}{2}} - \frac{E_0}{RT} - \frac{nE_1}{RT} \quad (15)$$

$$\frac{\partial\ln\left(\frac{dn}{dt}\right)}{\partial n} = -\frac{E_1}{RT} = -b \quad (16)$$

and

$$\frac{\partial\ln\left(\frac{dn}{dt}\right)}{\partial T} = \frac{E_0}{RT^2}+\frac{nE_1}{RT^2} = \frac{1}{RT^2}[E_0+E_1n]. \quad (17)$$

Thus, from equations (16) and (17), we can calculate values for E_0 and E_1, and the resulting values of ΔE^* above $\theta = 1$ are given by

$$\Delta E^* = 5{\cdot}76 \times 10^3 + 3{\cdot}65 \times 10^{12} (n - n_m) \text{ cal/mole.}$$

SUMMARY

Some details of the particle size of an evaporated silicon film as a function of its thickness have been obtained.

The slow oxygen uptake in the 1–2 monolayer region has been investigated. The rates are proportional to the square root of the oxygen pressure (indicating dissociation of the oxygen molecule in the rate-determining reaction). Illumination does not affect the rate. Over the temperature range covered (273–323°K), the activation energy is given by

$$\Delta E^* = 5{\cdot}76 \times 10^3 + 3{\cdot}65 \times 10^{11} (n - n_m) \text{ cal/mole.}$$

The kinetics are described by an Elovich-type equation, and evacuation experiments indicate that some type of rearrangement of sorbed material is the slow step.

Acknowledgement—The author wishes to express his indebtedness to P. S. MEIGS, who carried out most of the experimental work reported here.

REFERENCES

1. GREEN M. and KAFALAS J. A. *Phys. Rev.* **98,** 1566 (1955).
2. ARCHER R. J. Communication to Electrochemical Society Meeting, Washington, May 1957.
3. LAW J. T. *J. Phys. Chem.* **61,** 1200 (1957).
4. ALPERT D. *J. Appl. Phys.* **24,** 860 (1953).
   BILLS D. G. and ALLEN F. G. *Rev. Sci. Instrum.* **26,** 654 (1955).
5. ALLEN F. G. Private communication.
6. HONIG R. E. *J. Chem. Phys.* **22,** 1610 (1954).
7. ALLEN J. A. and MITCHELL J. W. *Disc. Faraday Soc.* No. 8, 309 (1950).
8. PORTER A. S. and TOMPKINS F. C. *Proc. Roy. Soc.* A **217,** 529 (1953).
9. BEECK O., SMITH A. E., and WHEELER A. *Proc. Roy. Soc.* A **177,** 62 (1941).
10. GREEN M., KAFALAS J. A. and ROBINSON P. H. *Semiconductor Surface Physics* p.349. University of Pennsylvania Press (1957).
11. STEPHENS S. J. Private communication.
12. HASS G. *Kolloidzcshr.* **100,** 230 (1942).
13. SCHOFFER H. *Z. Phys.* **130,** 565 (1951).
14. STRAUMANIS M. E. and AKA E. J. *J. Appl. Phys.* **23,** 230 (1952).
15. LANDSBERG P. T. *J. Chem. Phys.* **23,** 1079 (1955).
16. LANYON A. H. and TRAPNELL B. M. W. *Proc. Roy. Soc.* A **227,** 387 (1955).
17. DILLON J. A. *Bull. Amer. Phys. Soc.* **1,** 53 (1956).
18. DILLON J. A. and FARNSWORTH H. E. *Phys. Rev.* **99,** 1643 (1955).
19. CABRERA N. and MOTT N. F. *Ref. Progr. Phys.* **12,** 163 (1948).

VOL. 4 1958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 101–110.

THEORY OF CYCLOTRON-RESONANCE ABSORPTION BY CONDUCTION ELECTRONS IN INDIUM ANTIMONIDE

RICHARD F. WALLIS

U.S. Naval Research Laboratory, Washington 25, D.C.

(*Received* 28 *June* 1957)

Abstract—A semi-classical theory of cyclotron resonance by conduction electrons in InSb is presented. It is assumed that the conduction band is a simple band with a minimum at $k = 0$, and spin interactions are neglected. Terms up to the fourth order in the expansion of E in powers of k about the band edge are taken into account in a computation of the energy levels and wave functions of an electron in a constant magnetic field, using a simplified form of the effective mass equation of KJELDAAS and KOHN in which warping of the conduction-band energy surfaces is neglected. An expression for the absorption coefficient has been obtained which is accurate to terms linear in the fourth-order energy coefficient. Collision broadening is incorporated, using a single relaxation time. The fourth-order coefficient in the energy results in a splitting of the cyclotron-resonance absorption band into a series of peaks, each broadened towards the low-frequency side. The frequencies of the maxima of these peaks are shifted towards lower values by an amount which is proportional to the magnetic field and to the fourth-order coefficient in the energy and increases with magnetic quantum number. Using the energy-band parameters computed by KANE, the theoretical frequency shift is larger than that observed in infra-red cyclotron-resonance experiments, but it is of the right order of magnitude.

INTRODUCTION

CYCLOTRON-resonance absorption by conduction electrons in indium antimonide at room temperature has recently been observed by BURSTEIN *et al.*[1] using infra-red radiation and constant magnetic fields up to 60,000G. Similar observations, using pulsed magnetic fields up to 300,000G, have been reported by KEYES *et al.*[2] KANE[3] has shown theoretically that the conduction band in indium antimonide is strongly non-parabolic. The apparent effective mass of conduction electrons and the corresponding cyclotron-resonance frequency should accordingly vary rather rapidly with energy. The room-temperature cyclotron-resonance experiments with large magnetic fields may be expected to be influenced much more strongly by the variation of effective mass than the microwave experiments of DRESSELHAUS *et al.*,[4] carried out at temperatures below 4°K with small magnetic fields. It is the purpose of the present paper to investigate theoretically the effect of the non-parabolic nature of the conduction band on the room-temperature cyclotron-resonance absorption of conduction electrons in indium antimonide in relatively large magnetic fields.

CALCULATION OF ENERGIES AND EIGENFUNCTIONS

The Hamiltonian for a conduction electron of charge $-e$ in the presence of a time-dependent external electromagnetic field and a constant external magnetic field may be written in the form

$$\mathscr{H} = [\mathbf{p}+(e/c)(\mathbf{A}_0+\mathbf{A}_1)]^2/2m+V(\mathbf{r}) \qquad (1)$$

where $V(\mathbf{r})$ is the periodic potential for conduction electrons in indium antimonide without external fields and $\mathbf{A}_0$ and $\mathbf{A}_1$ are the vector potentials for the constant magnetic field and time-dependent electromagnetic field, respectively. Spin-orbit interactions as well as spin interactions with the external fields will be neglected throughout this paper. If one assumes (as is customary in semi-classical radiation theory) that the quantity $e^2\mathbf{A}_1^2/2mc^2$ is negligibly small and that p commutes with $\mathbf{A}_1$, the Hamiltonian can be written as

$$\mathscr{H} = \mathscr{H}_0 + \mathscr{H}_1 \tag{2}$$

where
$$\mathscr{H}_0 = \boldsymbol{P}^2/2m + V(\boldsymbol{r}) \tag{2a}$$

$$\mathscr{H}_1 = (e/mc)\boldsymbol{A}_1\boldsymbol{P} \tag{2b}$$

$$\boldsymbol{P} = \boldsymbol{p} + \frac{e}{c}\boldsymbol{A}_0. \tag{2c}$$

The time-dependent term $\mathscr{H}_1$ will be treated by semi-classical perturbation theory following the development of SEITZ[5] with appropriate modifications.

The eigenfunctions and eigenvalues for the operator $\mathscr{H}_0$ will be obtained by using the effective mass formalism of LUTTINGER and KOHN[6] carried to the fourth order. It will be assumed that the conduction band in indium antimonide in the absence of external fields is a simple band with a minimum at $k = 0$ and that the corresponding Bloch function for $k = 0$ belongs to the totally symmetric representation Γ_1 of the cubic group.[7] Under these conditions and with all spin interactions neglected, the effective mass equation for a conduction electron in the presence of a constant external magnetic field H can be written, correct to the fourth order, in the form

$$\begin{aligned}[P^2/2m^* + \mathscr{E}_c^{xxxx}\boldsymbol{P}^4 + (3\,\mathscr{E}_c^{xxyy} - \mathscr{E}_c^{xxxx}) \times \\ \times (P_x^2P_y^2 + P_y^2P_x^2 + P_y^2P_z^2 + P_z^2P_y^2 + \\ + P_z^2P_x^2 + P_x^2P_z^2) + \\ + J_c]F_c(\boldsymbol{r}) = (E - \mathscr{E}_c)F_c(\boldsymbol{r}).\end{aligned} \tag{3}$$

where m^* is the effective mass determined by the curvature of the conduction band at $k = 0$, E is the energy eigenvalue, $\mathscr{E}_c$ is the conduction-band energy at $k = 0$ in the absence of an external magnetic field,

$$s = eH/\hbar c, \tag{3a}$$

$$\mathscr{E}_c^{\alpha\beta\gamma\delta} = [\sum_P E_c^{\alpha\beta\gamma\delta}]/N(\boldsymbol{P}), \tag{3b}$$

$$J_c = \hbar^4 s^2[3\,\mathscr{E}_c^{xxyy} - E_c^{xxyy} + E_c^{xyyx}]. \tag{3c}$$

The symbol $\sum_P$ denotes a sum over all permutations of α, β, γ, and δ, and $N(\boldsymbol{P})$ is the number of such permutations. The quantities $E_c^{\alpha\beta\gamma\delta}$ are defined by

$$\begin{aligned}E_c^{\alpha\beta\gamma\delta} = \hbar^{-3}m^{-4}\{ \sum_{j\neq c} \sum_{i,k} p_{ci}^{\alpha}p_{ij}^{\beta}p_{jk}^{\gamma}p_{kc}^{\delta}/\omega_{ci}\omega_{cj}\omega_{ck} - \tfrac{1}{2}[\sum_i p_{ci}^{\alpha}p_{ic}^{\beta}/\omega_{ci}^2][\sum_k p_{ck}^{\gamma}p_{kc}^{\delta}/\omega_{ck}] - \\ - \tfrac{1}{2}[\sum_i p_{ci}^{\alpha}p_{ic}^{\beta}/\omega_{ci}][\sum_k p_{ck}^{\gamma}p_{kc}^{\delta}/\omega_{ck}^2] - (\hbar m/4)\sum_i [p_{ci}^{\alpha}p_{ic}^{\beta}\delta^{\gamma\delta} - 2p_{ci}^{\alpha}p_{ic}^{\delta}\delta^{\beta\gamma} + p_{ci}^{\gamma}p_{ic}^{\delta}\delta^{\alpha\beta}]/\omega_{ci}^2\}.\end{aligned} \tag{3d}$$

where p_{ij}^{α} is the matrix element of the α–th component of momentum with respect to Bloch functions for bands i and j at $k = 0$ in the absence of an external magnetic field and $\hbar\omega_{ij}$ is the energy difference between these bands at $k = 0$. Quantities such as $p_{ij}^{\alpha}/\omega_{ij}$ are taken to be zero when $i = j$. Equation (3) is a slight modification of an equation first given by KJELDAAS and KOHN.[8]

The quantities $\hbar^4\mathscr{E}_c^{\alpha\beta\gamma\delta}$ are the fourth-order coefficients in the expansion of the conduction-band energy about $k = 0$ in the absence of an external magnetic field. The term in equation (3) with coefficient $3\mathscr{E}_c^{xxyy} - \mathscr{E}_c^{xxxx}$ arises from a warping of the conduction-band energy surfaces in the fourth order. If one considers only the valence and conduction bands and neglects all other bands, the quantity $3\mathscr{E}_c^{xxyy} - \mathscr{E}_c^{xxxx}$ is zero and the third term in equation (3) vanishes. The results of KANE[3] indicate that for energies $< 0{\cdot}5$ eV one can make a fairly accurate calculation of the conduction-band energy in the absence of an external magnetic field by considering only the interactions of valence and conduction bands. Since we shall be concerned with energies $< 0{\cdot}2$ eV, we shall omit the third term in equation (3) and confine ourselves to the solution of

$$[\boldsymbol{P}^2/2m^* + \mathscr{E}_c^{xxxx}\boldsymbol{P}^4]F_c(\boldsymbol{r}) = (E - \mathscr{E}_c - J_c)F_c(\boldsymbol{r}). \tag{4}$$

The sign of $\mathscr{E}_c^{xxxx}$ given by KANE's theory is negative and will be considered as negative throughout this paper. A negative sign is consistent with the increase of effective mass with magnetic field observed in infra-red cyclotron-resonance experiments.[2]

VOL. 4 1958

To solve equation (4) we choose the gauge

$$\boldsymbol{A}_0 = (-Hy, 0, 0) \tag{5}$$

which specifies a constant magnetic field in the z-direction. The eigenfunctions for equation (4) are given by

$$F_c(\boldsymbol{r}) = s^{\frac{1}{2}}(L_xL_z)^{-\frac{1}{2}}Y_l(t)\exp[i(k_zx+k_zz)] \tag{6}$$

where

$$t = s^{\frac{1}{2}}[y-(k_x/s)], \tag{6a}$$

$$Y_l(t) = (2^ll!\sqrt{\pi})^{-\frac{1}{2}}H_l(t)\exp(-t^2/2), \tag{6b}$$

$H_l(t)$ is a Hermite polynomial of degree l, and L_x and L_z are the dimensions of the sample in the x- and z-directions. The functions $F_c(\boldsymbol{r})$ are normalized so that

$$\int_{-L_x/2}^{L_x/2}\int_{-L_z/2}^{L_z/2}\int_{-\infty}^{\infty} F_c^*(\boldsymbol{r})F_c(\boldsymbol{r})\,dxdydz = 1. \tag{6c}$$

If periodic boundary conditions are employed in the x- and z-directions, the quantum numbers k_x and k_z are given by

$$k_x = 2\pi n_x/L_x, \tag{6d}$$

$$k_z = 2\pi n_z/L_z, \tag{6e}$$

where n_x and n_z are integers.

The energy eigenvalues for equation (4) are given by

$$E(k_z,l) = \mathscr{E}_c+J_c+\hbar^2k_z^2/2m^*+\hbar\omega_c(l+\tfrac{1}{2})+ \\ +\hbar^4\,\mathscr{E}_c^{xxxx}[k_z^4+4s(l+\tfrac{1}{2})k_z^2+4s^2(l+\tfrac{1}{2})^2] \tag{7}$$

where

$$\omega_c = eH/m^*c. \tag{7a}$$

A result similar to equation (7) has been obtained previously.[9] The constant J_c in equation (7) is related to the "antisymmetric" constant discussed by KJELDAAS and KOHN.[8] This constant has no effect on cyclotron resonance, but may be of importance in the theory of the interband magneto-optic effect.

In Fig. 1 the energy specified by equation (7) is plotted schematically as a function of k_z. If the fourth-order terms were zero, the curves for various values of l would be parabolas which could be superposed by suitable translations along the vertical axis. Since cyclotron resonance involves vertical transitions from a magnetic level with quantum number l to the next higher magnetic level with quantum number $l+1$, the absorption

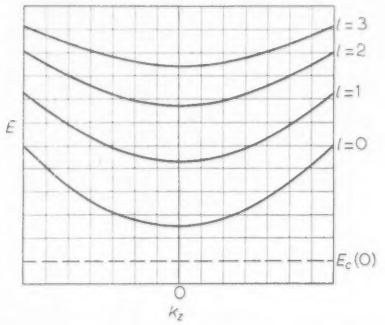


FIG. 1. Energy as a function of propagation constant k_z for various values of the magnetic quantum number l.

spectrum in the absence of fourth-order terms is a single peak. Inclusion of fourth-order terms with the sign of $\mathscr{E}_c^{xxxx}$ taken to be negative has two gross effects on the energy-level diagram. First, the separation between adjacent magnetic levels at $k_z = 0$ decreases as l increases. Second, the separation between adjacent magnetic levels decreases as the magnitude of k_z increases. The cyclotron-resonance absorption spectrum then should consist of a number of peaks each broadened to the low-energy side.

Within the framework of approximations mentioned earlier, equation (7) specifies to the fourth-order those energy eigenvalues of the Hamiltonian $\mathscr{H}_0$ which correspond to conduction electrons. The corresponding eigenfunctions of $\mathscr{H}_0$ can be found by using the methods of LUTTINGER and KOHN[6] and are given correct to the fourth order by

$$\Psi = N\Big\{F_c(\boldsymbol{r})u_c(\boldsymbol{r})-\sum_i\Big[\frac{\boldsymbol{p}_{ic\alpha}}{m\hbar\omega_{ic}}\cdot\boldsymbol{P}_\alpha F_c(\boldsymbol{r})\Big]u_i(\boldsymbol{r})- \\ -\frac{1}{2}\sum_i\Big[\frac{\boldsymbol{p}_{ci^\alpha}\boldsymbol{p}_{ic^\beta}}{m^2\hbar^2\omega_{ci}^2}\cdot\boldsymbol{P}_\alpha\boldsymbol{P}_\beta F_c(\boldsymbol{r})\Big]u_c(\boldsymbol{r})+$$

$$+\sum_{i\neq c}\sum_{j}\left[\frac{\boldsymbol{p}_{ij}{}^{\alpha}\boldsymbol{p}_{jc}{}^{\beta}}{m^2\hbar^2\omega_{ci}\omega_{cj}}\cdot\boldsymbol{P}_{\alpha}\boldsymbol{P}_{\beta}F_c(\boldsymbol{r})\right]u_i(\boldsymbol{r})+$$

$$+\sum_{i\neq c}\sum_{j\neq c}\sum_{k}\left[\frac{\boldsymbol{p}_{ij}{}^{\alpha}\boldsymbol{p}_{jk}{}^{\beta}\boldsymbol{p}_{kc}}{m^3\hbar^3\omega_{ci}\omega_{cj}\omega_{ck}}\cdot\boldsymbol{P}_{\alpha}\boldsymbol{P}_{\beta}\boldsymbol{P}_{\gamma}F_c(\boldsymbol{r})\right]u_i(\boldsymbol{r})-$$

$$-\sum_{i\neq c}\sum_{j\neq c}\left[\frac{1}{m^3\hbar^3}\left(\frac{1}{2}\cdot\frac{1}{\omega_{ci}\omega_{cj}{}^2}+\frac{1}{\omega_{c_i}{}^2\omega_{c_j}}\right)\times\right.$$

$$\left.\times\,\boldsymbol{p}_{ic}{}^{\alpha}\boldsymbol{p}_{cj}{}^{\beta}\boldsymbol{p}_{jc}{}^{\gamma}\boldsymbol{P}_{\alpha}\boldsymbol{P}_{\beta}\boldsymbol{P}_{\gamma}F_c(\boldsymbol{r})\right]u_i(\boldsymbol{r})\Bigg\} \tag{8}$$

where $u_i(\boldsymbol{r})$ is the Bloch function for the ith band at $k=0$. The normalization constant N is $(2\pi)^{3/2}$ if the functions $u_i(\boldsymbol{r})$ are normalized so that

$$\int_{\text{cell}} u_i{}^*(\boldsymbol{r})u_i(\boldsymbol{r})\,d\boldsymbol{r}=\Omega/(2\pi)^3 \tag{8a}$$

where Ω is the volume of a unit cell.

CALCULATION OF THE ABSORPTION COEFFICIENT

The semi-classical theory of the absorption of radiation[5] can be generalized in a straightforward manner to the case in which the absorber is an electron in a solid in the presence of an external magnetic field. The absorption coefficient for radiation of circular frequency ω associated with a transition between quantum states i and j can be written as

$$k(\omega)=(4\pi^2ne^2/\epsilon cm^2\omega)|\boldsymbol{\varkappa}\cdot\boldsymbol{P}_{ji}|^2\times \times\delta(E_i-E_j+\hbar\omega) \tag{9}$$

$$=(4\pi^2ne^2/\epsilon cm^2\omega_{ji})|\boldsymbol{\varkappa}\cdot\boldsymbol{P}_{ji}|^2\times \times\delta(E_i-E_j+\hbar\omega) \tag{9a}$$

where n and ϵ are the index of refraction and the real part of the dielectric constant of the solid, $\boldsymbol{\varkappa}$ is the polarization vector for the radiation,

$$\omega_{ji}=(E_j-E_i)/\hbar \tag{9b}$$

and

$$\boldsymbol{P}_{ji}=\int\Psi_j{}^*\boldsymbol{P}\Psi_i\,d\boldsymbol{r}. \tag{9c}$$

The wave functions Ψ_i and Ψ_j both have the form given by equation (8). The function $F_c(\boldsymbol{r})$ appearing in Ψ_i has quantum numbers k_x, k_z, and l and will be denoted by $F_{ci}(\boldsymbol{r})$, while the function $F_c(\boldsymbol{r})$ appearing in Ψ_j has a different set of quantum numbers, say k_x', k_z' and l', and will be denoted by $F_{cj}(\boldsymbol{r})$. If one substitutes equation (8) into equation (9c) and simplifies the various integrals by using the methods of Luttinger and Kohn,[6] one finds that

$$\boldsymbol{P}_{ji}{}^{\alpha}=(\boldsymbol{P}_{ji}{}^{\alpha})_1+(\boldsymbol{P}_{ji}{}^{\alpha})_2+(\boldsymbol{P}_{ji}{}^{\alpha})_3 \tag{10}$$

$$(\boldsymbol{P}_{ji}{}^{\alpha})_1=\int F_{cj}{}^*(\boldsymbol{r})\boldsymbol{P}_{\alpha}F_{ci}(\boldsymbol{r})\,d\boldsymbol{r} \tag{10a}$$

$$(\boldsymbol{P}_{ji}{}^{\alpha})_2=\left(\frac{m}{m^*}-1\right)\times \times\int F_{cj}{}^*(\boldsymbol{r})\boldsymbol{P}_{\alpha}F_{ci}(\boldsymbol{r})\,d\boldsymbol{r} \tag{10b}$$

$$(\boldsymbol{P}_{ji}{}^{\alpha})_3=4m\,\mathscr{E}^{\alpha\beta\gamma\delta}\int F_{cj}{}^*(\boldsymbol{r})\times \times\boldsymbol{P}_{\beta}\boldsymbol{P}_{\gamma}\boldsymbol{P}_{\delta}F_{ci}(\boldsymbol{r})\,d\boldsymbol{r} \tag{10c}$$

where $\boldsymbol{P}_{ji}{}^{\alpha}$ is the component of $\boldsymbol{P}_{ji}$ in the α-direction. The term $(\boldsymbol{P}_{ji}{}^{\alpha})_1$ is a first-order term and represents the matrix element of $\boldsymbol{P}_{\alpha}$ with respect to the first terms in Ψ_i and Ψ_j as given in equation (8). The term $(\boldsymbol{P}_{ji\alpha})_2$ is a second-order term which arises from the matrix elements of $\boldsymbol{P}_{\alpha}$ with respect to the first term in Ψ_i and the second term in Ψ_j and the reverse. Since the ratio m/m^* for conduction

VOL. 4 1958

electrons in InSb is about 77, the second-order term $(\boldsymbol{P}_{ji}{}^{\alpha})_2$ is much more important than the first-order term $(\boldsymbol{P}_{ji}{}^{\alpha})_1$.

The components $\boldsymbol{P}_{ji}{}^{\alpha}$ can be evaluated using equation (6). If one neglects the warping terms in the conduction band, i.e. assumes $\mathscr{E}_c{}^{xxxx} = 3\mathscr{E}_c{}^{xxyy}$, the results are

$$\boldsymbol{P}_{ji}{}^{x} = \hbar s^{\frac{1}{2}}\left\{\frac{m}{m^*}\left[\sqrt{\frac{l}{2}}\,\delta_{l'l-1}+\sqrt{\frac{l+1}{2}}\,\delta_{l'l+1}\right]+ +4m\,\mathscr{E}_c{}^{xxxx}\left[(2m^*E_2-\hbar^2 s)\sqrt{\frac{l}{2}}\,\delta_{l'l-1}+ +(2m^*E_2+\hbar^2 s)\sqrt{\frac{l+1}{2}}\,\delta_{l'l+1}\right]\right\}\delta_{k_x k'_x}\delta_{k_z k'_z} \tag{11a}$$

$$\boldsymbol{P}_{ji}{}^{y} = i\hbar s^{\frac{1}{2}}\left\{\frac{m}{m^*}\left[-\sqrt{\frac{l}{2}}\,\delta_{l'l-1}+\sqrt{\frac{l+1}{2}}\,\delta_{l'l+1}\right]+ +4m\,\mathscr{E}_c{}^{xxxx}\left[-(2m^*E_2-\hbar^2 s)\sqrt{\frac{l}{2}}\,\delta_{l'l-1}+ +(2m^*E_2+\hbar^2 s)\sqrt{\frac{l+1}{2}}\,\delta_{l'l+1}\right]\right\}\delta_{k_x k'_x}\delta_{k_z k'_z} \tag{11b}$$

where

$$E_2 = \frac{\hbar^2 k_z{}^2}{2m^*}+\hbar\omega_c(l+\tfrac{1}{2}). \tag{11c}$$

The selection rules specified by equation (11) for cyclotron resonance are $l' = l\pm 1$, $k_x = k_x'$, $k_z{}' = k_z$. If the warping of the conduction band were not neglected, it appears that the terms $(\boldsymbol{P}_{ji}{}^{x})_3$ and $(\boldsymbol{P}_{ji}{}^{y})_3$ would be non-vanishing for $l' = l\pm 3$ and third harmonics would be allowed. For cyclotron resonance the z-component of $\boldsymbol{P}_{ji}$ does not give transitions between quantum states of different energies and will not be considered further.

The expression for the absorption coefficient given by equation (9a) neglects broadening. We shall assume that the absorption due to a transition between two given quantum states is broadened according to a Lorentz shape function having a relaxation time τ which is independent of the particular transition involved. The broadening is achieved formally by replacing $\hbar\delta(E_i-E_j+\hbar\omega)$ in equation (9a) by a Lorentz shape function $\mathscr{S}(\omega_{ji}, \omega)$ given by

$$\mathscr{S}(\omega_{ji}, \omega) = \frac{2\tau}{\tau^2(\omega_{ji}-\omega)^2+1}. \tag{12}$$

The form of $\mathscr{S}(\omega_{ji}, \omega)$ specified by equation (12) is appropriate if $\omega_{ji} \gg 1/\tau$, so that in the region of significant absorption $\omega \simeq \omega_{ji}$.

The total absorption coefficient at circular frequency ω is obtained by summing the contributions from all transitions. If equations (11a) and (11b) for $\boldsymbol{P}_{ji}{}^{x}$ and $\boldsymbol{P}_{ji}{}^{y}$ are employed, the total absorption coefficient, including broadening for radiation plane polarized in either the x- or y-direction, can be written as

$$k(\omega) = (4\pi^2 ne^2/\epsilon\hbar cm^2)\sum_{k_x}\sum_{k_z}\sum_{l}[f(k_x, k_z, l)-f(k_x, k_z, l+1)]\times (1/\omega_{ji})\cdot\hbar^2 s\left(\frac{m}{m^*}\right)^2\left(\frac{l+1}{2}\right)[1+8m^*\,\mathscr{E}_c{}^{xxxx}(2m^*E_2+\hbar^2 s)\times \mathscr{S}(\omega_{ji}, \omega) \tag{13}$$

where $f(k_x, k_z, l)$ is the number of electrons per unit volume having quantum numbers k_x, k_z, l. The term involving $f(k_x, k_z, l+1)$ is a correction for induced emission. It will be assumed that the distribution function f is a Maxwell–Boltzmann distribution,

$$f(k_x, k_z, l) = A \exp[-E(k_z, l)/kT]. \quad (14)$$

This assumption is not completely justified, because we shall consider concentrations of conduction electrons which lead to a value of the order of unity for the degeneracy parameter $N_0h^3/2(2\pi m^*kT)^{3/2}$. The use of the Fermi–Dirac distribution, however, leads to considerable mathematical difficulty and has accordingly not been attempted in the present paper.

The constant A in equation (14) is determined by the equation

$$2\sum_{k_x}\sum_{k_z}\sum_{l} f(k_x, k_z, l) = N_0 \quad (15)$$

where N_0 is the concentration of conduction electrons and the factor 2 takes into account the two-fold spin degeneracy of each state. The values of k_x range from zero to a maximum value[11] determined by the requirement that the center of the electron orbit be within the crystal. The total number of permitted values of k_x per unit volume is $(eH/hc)(L_xL_y/V)$, where L_x and L_y are the dimensions of the crystal in the x- and y-directions and V is the volume of the crystal. The left-hand side of equation (15) can be evaluated by replacing the sum over k_z by an integral and using equations (7) and (14). The result correct to the first power of $\mathscr{E}_c{}^{xxxx}$ can be written as

$$A = 4\pi^2N_0s^{-3/2}\{1-\mu[(3/2\beta)+ + 2(1+2S_1)+2\beta(1+4S_1+4S_2)]\}\times \times(\beta/\pi)^{\frac{1}{2}}\sinh\beta \quad (16)$$

where

$$\mu = m^{*2}\hbar\omega_c|\mathscr{E}_c{}^{xxxx}| \quad (16a)$$

$$\beta = \hbar\omega_c/2kT. \quad (16b)$$

$$S_1 = [\exp(-\beta)]/(2\sinh\beta) \quad (16c)$$

$$S_2 = S_1\coth\beta. \quad (16d)$$

It is convenient to express the total absorption coefficient $k(\omega)$ in terms of the conductivity $\sigma(\omega)$ by means of the relation

$$k(\omega) = (4\pi n/\epsilon c)\sigma(\omega). \quad (17)$$

By making use of equations (7) and (11c) and replacing the sum over k_z in equation (13) by an integral, one can write the conductivity in the form

$$\sigma(\omega) = \sum_{l=0}^{\infty}\sigma_l(\omega) \quad (18)$$

$$\sigma_l(\omega) = (\alpha c/8\pi^2)As^{\frac{1}{2}}\times (1+3\mu/2\beta)(\omega_l/\omega_c)(l+1)\times \{e^{-a(l)}K_l(\omega)-e^{-a(l+1)}K_l'(\omega)\} \quad (18a)$$

where α is the fine-structure constant,

$$\omega_l = \omega_c[1-8\mu(l+1)-2\mu/\beta], \quad (18b)$$

$$a(l) = 2(l+\tfrac{1}{2})[1-4\mu(l+\tfrac{1}{2})]\beta, \quad (18c)$$

$$K_l(\omega) = \omega_c\int_{-\infty}^{\infty} dx\,\mathscr{S}_l(\omega_{jl},\omega)\times \exp\{-b(l)x^2\}, \quad (18d)$$

$$K_l'(\omega) = \omega_c\int_{-\infty}^{\infty} dx\,\mathscr{S}_l(\omega_{jl},\omega)\times \exp\{-b(l+1)x^2\}, \quad (18e)$$

$$b(l) = [1-8\mu(l+\tfrac{1}{2})]\beta, \quad (18f)$$

$x = k_z/s^{\frac{1}{2}}$, and $\mathscr{S}_l(\omega_{jl},\omega)$ is obtained from $\mathscr{S}(\omega_{jl},\omega)$ by substituting the value of ω_{jl} corresponding to a transition from a state with quantum numbers l, k_x, k_z to a state with quantum numbers $l+1, k_x, k_z$. The evaluation of the integrals $K_l(\omega)$ and $K_l'(\omega)$ is discussed in the Appendix. The quantity $\sigma_l(\omega)$ is the contribution to the conductivity associated with transitions in which l changes to $l+1$ and is given correct to linear terms in the fourth-order constant μ by equation (18a). To within the accuracy of the treatment, the maximum value of $\sigma_l(\omega)$ occurs at the frequency ω_l.

In Fig. 2 the conductivities $\sigma_0(\omega)$ and $\sigma_1(\omega)$ associated with the 0–1 and 1–2 transitions are plotted as functions of frequency for a conduction-electron concentration of 5×10^{16}cm^{-3}, magnetic

field of 80,000G, and a relaxation time given by $\omega_c\tau = 10$. The effective mass used is that obtained by DRESSELHAUS *et al.*[4] The value of the fourth-order constant μ was chosen arbitrarily so that the calculations would show clearly the effects of the

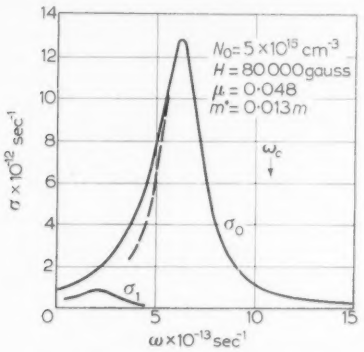


FIG. 2. Conductivity as a function of frequency. The value of $\omega_c\tau$ is 10.

fourth-order term and yet be within the range of validity of the theory. If the fourth-order term were omitted, the maxima of both $\sigma_0(\omega)$ and $\sigma_1(\omega)$ would occur at the frequency ω_c indicated by the arrow in Fig. 2. The fourth-order term causes $\sigma_0(\omega)$ to have its maximum to the low-frequency side of ω_c and $\sigma_1(\omega)$ to have its maximum to the low-frequency side of both ω_c and the maximum of $\sigma_0(\omega)$. One can introduce apparent effective masses m_0^* and m_1^* defined by the equation

$$\omega_l = eH/m_l^*c \tag{19}$$

with $l = 0$ and $l = 1$. Using equation (18*b*), one can obtain the relation

$$1/m_l^* = (1/m^*)[1-8\mu(l+1)-2\mu/\beta] \tag{20a}$$

or since we are assuming that $\mu \ll 1$,

$$m_l^* = m^*[1+8\mu(l+1)+2\mu/\beta]. \tag{20b}$$

Since μ is proportional to the magnetic field, the apparent effective mass m_l^* increases linearly with magnetic field. From the frequencies at which $\sigma_0(\omega)$ and $\sigma_1(\omega)$ have their maximum values, one can compute m_0^* and m_1^* by using equation (19). The values obtained from the cases given in Fig. 2 are $m_0^* = 0{\cdot}022$ m and $m_1^* = 0{\cdot}06$ m.

The broadening introduced to the low-frequency side of $\sigma_0(\omega)$ and $\sigma_1(\omega)$ by the fourth-order term is evident in Fig. 2. The broadening is emphasized for $\sigma_0(\omega)$ by the dotted line which is the reflection of the high-frequency side about the maximum.

The absorption coefficient $k(\omega)$ can be computed from the conductivity by using equation (17), provided one knows the index of refraction n and dielectric constant ϵ as functions of frequency. For a material having of the order of 10^{16} conduction electrons with very high mobility, one should investigate the extent to which n and ϵ vary with frequency. The quantities n and ϵ are related to the conductivity σ and polarizability α_p by the equations

$$\epsilon = 1+4\pi\alpha \tag{21a}$$

$$n^2 = \tfrac{1}{2}\{\epsilon+[\epsilon^2+16\pi^2(\sigma/\omega)^2]^{\frac{1}{2}}\}. \tag{21b}$$

The conductivity σ is given by equation (18). We shall assume that α_p is the sum of two terms

$$\alpha_p = \alpha_{p0}+\alpha_p(\omega) \tag{22}$$

where α_{p0} is the polarizability of InSb far from resonance and $\alpha_p(\omega)$ is the polarizability contribution of the conduction electrons in the region of absorption. The quantity $\alpha_p(\omega)$ can be calculated by using the semi-classical theory by a suitable modification of the treatment of SEITZ.[5] The contribution of a given quantum state i to the polarizability tensor can be written as

$$\boldsymbol{\alpha}_i(\omega) = \sum_{j\neq i} \frac{\mathcal{R}[\boldsymbol{M}_{ji}^*\boldsymbol{M}_{ji}]}{\hbar} \frac{2\omega_{ji}}{\omega_{ji}^2-\omega^2} \tag{23}$$

where

$$\boldsymbol{M}_{ji} = (e/im\omega_{ji})\boldsymbol{P}_{ji} \tag{23a}$$

and $\mathcal{R}$ signifies the real part. For a magnetic field in the z-direction and radiation plane polarized in the x- or y-direction, the polarizability contribution $\alpha_p(\omega)$ can be found by summing equation (23) over all states i and making use of equations (7), (9), and (11). The results can be written as

$$\alpha_p(\omega) = \tau\sum_{l=0}^{\infty} \frac{\omega_l-\omega}{\omega_l}\sigma_l(\omega) \tag{24}$$

where $\sigma_l(\omega)$ is given by equation (18*a*).

The index of refraction n and dielectric constant

ϵ were calculated as functions of frequency using equations (21) and (22) for the values of the parameters N_0, μ, H, m^*, and $\omega_c\tau$ considered in Fig. 2. The value of the quantity α_{p_0} was chosen so that $\epsilon = 16$ far from the region of absorption.

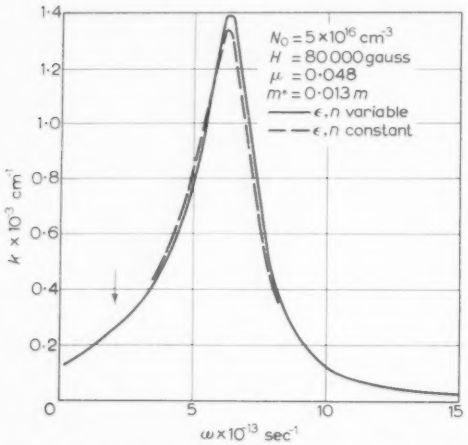


FIG. 3. Absorption coefficient as a function of frequency. The value of $\omega_c\tau$ is 10. The arrow indicates the maximum absorption due to the 1–2 transition.

The absorption coefficient was then calculated as a function of frequency in two different ways using equation (17). First, n and ϵ were taken to be independent of frequency with the values $n^2 = \epsilon = 16$. The result is indicated in Fig. 3 by the broken curve. Second, n and ϵ were taken to be frequency-dependent with the values calculated from equations (21) and (22). The result for $k(\omega)$ is shown by the solid curve in Fig. 3. It is evident that the frequency-dependence of n and ϵ produces a shift of the absorption peak toward higher frequencies which is rather small for the values of N_0 and m^* considered.

The arrow in Fig. 3 indicates the position of the maximum absorption due to the 1–2 transition. It is rather doubtful whether the peak due to this transition can be observed experimentally unless one can obtain samples having quite large values of τ. The small value of the intensity for the 1–2 transition compared to that for the 0–1 transition is a consequence of a number of factors which include the smaller population of the $l = 1$ level compared to the $l = 0$ level, the presence of the factor ω_l in the expression for $\sigma_l(\omega)$, and the relatively larger correction for induced emission for the 1–2 transition.

DISCUSSION

In order to make a detailed comparison between the present theoretical results and those of experiment, it is desirable that one should have experimental values for the absorption coefficient as a function of frequency at constant magnetic field. The available experimental data give the absorption coefficient as a function of magnetic field at a fixed frequency. Such data are difficult to compare with the theory because the concentration of conduction electrons depends on the magnetic field and because the fourth-order constant μ is proportional to the magnetic field. The degree of validity of the theory, which depends on the value of μ, will accordingly vary with the field. For these reasons no detailed comparison with experiment is made in the present paper.

It is possible, however, to obtain an estimate of the fourth-order constant from the linear dependence of the apparent effective mass on magnetic field observed by KEYES *et al.*[2] for fields up to 210,000G. According to equation (20*b*), the apparent effective mass associated with the 0–1 transition should increase linearly with magnetic field with a slope given by $8\mu m^*/H$. Using the slope of the experimental data given by KEYES *et al.*, one finds that for $H = 80{,}000$G and $m^* = 0{\cdot}013\ m$, $\mu = 0{\cdot}06$.

A theoretical estimate of μ can be obtained from KANE's[3] theory of the energy-band structure of InSb. If one considers only the interactions between conduction and valence bands, KANE's theory leads to the following expression for the fourth-order constant $\mathscr{E}_c{}^{xxxx}$.

$$\mathscr{E}_c{}^{xxxx} = -[(m/m^*)-1]^2/4E_G m^2 \qquad (25)$$

where E_G is the separation of valence and conduction bands at $k = 0$. Taking $E_G = 0{\cdot}18$ eV,[12] $m^* = 0{\cdot}013$ m, and $H = 80{,}000$G and using equation (16*a*) one obtains $\mu = 0{\cdot}096$, which is in order-of-magnitude agreement with the value obtained from the data of KEYES *et al.*[2] It appears from KANE's work that the agreement will be lessened rather than improved by including interactions of valence and conduction bands with higher bands in estimating the fourth-order constant. A possible reason for the discrepancy is that

VOL. 4 1958

terms beyond the fourth order should be included in the effective mass formalism for calculating the energy levels and wave functions in the presence of a magnetic field.

DRESSELHAUS[7] has made a calculation of the energy-band structure of zinc blende structures including spin-orbit coupling. His results indicate that third-order terms remove the two-fold spin degeneracy of the conduction band in InSb at points other than $k = 0$ for most directions in k-space. The third-order terms might lead to observable effects in cyclotron resonance which should have a directional dependence. The author is not aware of any directional effects in the experiments on conduction-electron cyclotron resonance in InSb and has accordingly not included the third-order terms in the present theory.

The Maxwell–Boltzmann distribution employed in the present paper weights the various transitions with a factor which decreases exponentially as the value of k_z^2 for the initial state increases. The use of the Fermi–Dirac distribution would tend to give a more uniform weight to transitions involving initial states with values of k_z^2 up to a maximum value determined by the Fermi level. The broadening of the cyclotron-resonance absorption curves at the low-frequency side may then be expected to be greater than that given by the Maxwell–Boltzmann distribution.

The effective mass formalism employed in this paper is that appropriate for non-degenerate bands without spin-orbit coupling. If one wishes to take explicit account of spin-orbit coupling and of the two possible spin states for a conduction electron, one should use an effective mass formalism for degenerate bands similar to that given by LUTTINGER[15] for the valence bands of germanium. A possible consequence of such a treatment is that each absorption peak given by the simpler theory will be split into two or more peaks which may or may not be resolved. A further consequence may be that the apparent effective mass associated with a given peak may not have the dependence on magnetic field given by equation (20*b*).

Acknowledgements—I wish to thank ELIAS BURSTEIN for providing the stimulation to undertake this problem and for many helpful discussions. I am grateful to H. J. BOWLDEN, E. O. KANE, T. KJELDAAS, W. KOHN, M. LAX and G. S. PICUS for helpful discussions.

Appendix: Evaluation of the Integrals $K_l(\omega)$ and $K_l'(\omega)$.

Using equations (12) and (18*d*), one can write $K_l(\omega)$ in the form

$$K_l(\omega) = (I_1 - I_2)/4\mu i \tag{26}$$

where

$$I_1 = \int_{-\infty}^{\infty} dx(x^2 - c_1 - ic_2)^{-1} \times \exp[-b(l)x^2] \tag{26a}$$

$$I_2 = \int_{-\infty}^{\infty} dx(x^2 - c_1 + ic_2)^{-1} \times \exp[-b(l)x^2] \tag{26b}$$

$$c_1 = [1 - 8\mu(l+1) - (\omega/\omega_c)]/4\mu \tag{26c}$$

and

$$c_2 = \frac{1}{4\mu\omega_c\tau}. \tag{26d}$$

It is convenient to consider separately the cases $c_1 > 0$ and $c_1 < 0$. For $c_1 > 0$ one can use the method of partial fractions to reduce I_1 and I_2 to the forms

$$I_1 = [i/(\gamma_1 + i\gamma_2)]\chi(x, \eta) \tag{27a}$$

$$I_2 = -[i/(\gamma_1 - i\gamma_2)]\chi(-x, \eta) \tag{27b}$$

where

$$\gamma_1 = (c_1^2 + c_2^2)^{\frac{1}{4}} \cos(\phi/2) \tag{27c}$$

$$\gamma_2 = (c_1^2 + c_2^2)^{\frac{1}{4}} \sin(\phi/2) \tag{27d}$$

$$\phi = \arctan(|c_2|/|c_1|) \tag{27e}$$

$$x = \gamma_1/\gamma_2 \tag{27f}$$

$$\eta = 1/\gamma_2\sqrt{[b(l)]} \tag{27g}$$

and $\chi(x, \eta)$ is a function which has been discussed by BORN[13] in connection with the theory of Doppler broadened spectral lines. Using equation (27), equation (26) can be rewritten as

$$K_l(\omega) = [\pi/b(l)]^{\frac{1}{2}}[2u(\gamma_1^2+\gamma_2^2)]^{-1} \times [x\psi(x,\eta)+\phi(x,\eta)] \tag{28}$$

where $\psi(x,\eta)$ and $\phi(x,\eta)$ are related to $\chi(x,\eta)$ by

$$\chi(x,\eta) = [\psi(x,\eta)+i\phi(x,\eta)\eta\sqrt{\pi}]. \tag{28a}$$

Tables of $\psi(x,\eta)$ for various values of x and η are available.[14] Values of $\phi(x,\eta)$ can be obtained from values of $\psi(x,\eta)$ by means of the relation

$$\phi(x,\eta) = x\psi(x,\eta)+(\eta^2/2)\frac{\partial\psi(x,\eta)}{\partial x}. \tag{28b}$$

For $c_1 < 0$ one can show that

$$K_l(\omega) = [\pi/b(l)]^{\frac{1}{2}}[2\mu(\gamma_1^2+\gamma_2^2)]^{-1} \times [x'\psi(x',\eta')+\phi(x',\eta')] \tag{29}$$

where

$$x' = \gamma_2/\gamma_1 \tag{29a}$$

$$\eta' = 1/\gamma_1\sqrt{[b(l)]}. \tag{29b}$$

Formulas for $K_l'(\omega)$ can be obtained from those for $K_l(\omega)$ by replacing $b(l)$ by $b(l+1)$.

REFERENCES

1. Burstein E., Picus G. S. and Gebbie H. A. *Phys. Rev.* **103,** 825 (1956).
2. Keyes, R. J., Zwerdling S., Foner S., Kolm H. H. and Lax B. *Phys. Rev.* **104,** 1804 (1956).
3. Kane, E. O. *J. Phys. Chem. Solids* **1,** 249 (1956).
4. Dresselhaus G., Kip A. F., Kittel C. and Wagoner G. *Phys. Rev.* **98,** 556 (1955).
5. Seitz F. *Modern Theory of Solids* p. 215. McGraw-Hill, New York (1940).
6. Luttinger J. M. and Kohn W. *Phys. Rev.* **97,** 869 (1955).
7. Dresselhaus G. *Phys. Rev.* **100,** 580 (1955).
8. Kjeldaas T. and Kohn W. *Phys. Rev.* **105,** 806 (1957).
9. Zwerdling S., Keyes R. J., Foner S., Kolm H. H. and Lax B. *Phys. Rev.* **104,** 1807 (1956).
10. Breckenridge R. G., Blunt R. F., Hosler W. R., Frederikse H. P. R., Becker J. H. and Oshinsky W. *Phys. Rev.* **96,** 571 (1954).
11. Seitz F. *Modern Theory of Solids* p.584. McGraw-Hill, New York (1940).
12. Fan H. Y. and Gobeli G. W. *Bull. Amer. Phys. Soc.* **1,** 111 (1955).
13. Born M. *Optik* p. 482. Springer, Berlin (1933).
14. *U.S. Atomic Energy Publ.* **WAPD-SR-506** (Vols. I and II) (1954).
15. Luttinger J. M. *Phys. Rev.* **102,** 1030 (1956).

VOL. 4 1958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 111–117.

SUR L'ANNIHILATION DU POSITON DANS LES MÉTAUX

E. DANIEL et J. FRIEDEL

Physique des Solides, Sorbonne, Paris

(*Received* 13 *June* 1957)

Résumé—On montre que les temps de vie observés expérimentalement dans les métaux pour le positon peuvent s'expliquer par la polarisation des électrons de valence, sans formation de positronium. En admettant une interaction coulombienne avec écran entre positon et électrons, on trouve un temps de vie théorique de $0{,}7 \cdot 10^{-10}$ sec dans le cuivre, un peu inférieur aux valeurs expérimentales de $1{,}2 \pm 0{,}5 \cdot 10^{-10}$ sec et $2{,}7 \pm 0{,}3 \cdot 10^{-10}$ sec données respectivement par BELL et GRAHAM d'une part et GERHOLM d'autre part. On montre également que l'approximation grossière du puits carré de potentiel permet d'obtenir l'ordre de grandeur des résultats expérimentaux. On calcule aussi l'influence de l'effet Auger sur la corrélation angulaire des photons γ d'annihilation.

Abstract—It is shown that the positron lifetime observed experimentally in metals can be explained on the basis of the polarisation of the electron gas in the valency band, without forming a positronium. With a screened coulombian positron-electron interaction a lifetime of $0{\cdot}7 \times 10^{-10}$ sec is found, somewhat smaller than the experimental values $1{\cdot}2 \pm 0{\cdot}5 \times 10^{-10}$ sec and $2{\cdot}7 \pm 0{\cdot}3 \times 10^{-10}$ sec given respectively by BELL and GRAHAM and by GERHOLM. The order of magnitude for the experimental results can also be obtained by the cruder approximation of a square-well potential. The dependence on Auger effect of the γ-rays angular correlation is also computed.

1. INTRODUCTION

LE résultat de la théorie des collisions pour le système électron-positon ne s'applique pas directement à l'annihilation d'un positon dans un métal. On aurait,[1] pour un positon voyant la densité moyenne de la totalité* des électrons dans un métal de numéro atomique Z, contenant N atomes par centimètre cube, un taux d'annihilation :

$$\lambda \text{ sec}^{-1} = ZN\pi r_0^2 c \qquad (1)$$

où r_0 = rayon classique de l'électron et c = vitesse de la lumière. Ceci donne pour le temps de vie $\tau = \lambda^{-1}$ du positon des valeurs telles que $0{,}5 \cdot 10^{-10}$ sec dans le plomb et le cuivre, $1{,}7 \cdot 10^{-10}$ sec dans l'aluminium $5 \cdot 10^{-10}$ sec dans le sodium, variant beaucoup d'un métal à l'autre. Or les résultats des déterminations expérimentales dépendent assez peu du métal considéré : BELL et GRAHAM,[2] et DE BENEDETTI et RICHING,[3] donnent à peu près uniformément un temps de $1{,}5 \cdot 10^{-10}$ sec. Plus récemment, GERHOLM[4] et MINTON[5] (cf. aussi FERRELL[6]) ont trouvé des valeurs voisines de $2{,}5 \cdot 10^{-10}$ sec. Les incertitudes admises par les auteurs sur leurs résultats sont assez grandes, allant de 10 à 40 pour cent.

* C'est-à-dire Z électrons par atome.

2. EFFETS DES INTERACTIONS ÉLECTRONS-POSITON

Le premier effet des interactions des électrons du métal avec le positon est de ralentir celui-ci, jusqu'à le "thermaliser". LEE WHITING[7] a montré que ceci devait se faire dans un temps très court par rapport au temps de vie du positon dans le métal, de l'ordre de 3.10^{-12} sec. Au moment de son annihilation, le positon se trouve donc certainement loin des noyeaux atomiques, qui le repoussent fortement; il est par suite dans une région où il ne peut s'annihiler qu'avec un *électron de valence* du métal.

Comme il est à peu près au repos, il attire certainement les électrons de valence pour faire *écran* à sa charge positive. Le positon voit ainsi une densité électronique beaucoup plus forte que la densité moyenne des électrons de valence dans le métal pur. C'est cette densité électronique réelle au positon qu'il convient de substituer à ZN dans

l'équation (1). En fait, un temps de vie de $1{,}5 \cdot 10^{-10}$ sec nécessite une densité électronique au positon dix fois plus forte que la densité moyenne des électrons de valence dans un métal monovalent.

3. MÉTHODES UTILISÉES POUR LE CALCUL DU TEMPS DE VIE

Pour rendre compte de ses résultats expérimentaux, GERHOLM[4] suppose la formation d'un positronium (système lié électron-positon) dans le métal. En admettant que l'électron lié forme écran pour le positon, la densité électronique vue par celui-ci serait essentiellement la somme de la probabilité de présence de l'électron lié au positon, et de la densité moyenne des électrons de valence du métal. Ce modèle présente l'inconvénient de trop schématiser le phénomène de polarisation; d'autre part il n'est pas en accord avec les corrélations angulaires des photons γ émis lors de l'annihilation, si l'on admet que le positon est bien thermalisé. La formation d'un positronium dans le métal demeure douteuse; on peut expliquer au moins l'ordre de grandeur des temps de vie observés sans faire cette hypothèse.

Un calcul ne faisant intervenir que la polarisation des électrons de valence du métal a été développé par Ferrell.[6] Ce calcul, assez compliqué, repose sur certaines simplifications un peu discutables : il utilise une méthode de variation, qui exclut à priori la possibilité d'états liés et ne satisfait pas bien à la condition d'écran autour du positon. Le résultat du calcul du temps de vie, ajusté pour l'aluminium, n'est plus très bon quand on passe à d'autres métaux, par exemple le sodium.

Envisageant le problème d'un point de vue différent, nous faisons choix d'un potential raisonnable, à symétrie sphérique autour du positon et dépendant d'un paramètre. Nous déterminons ce paramètre en posant que le positon attire au total dans son voisinage une quantité de charge unité qui lui fait écran à grande distance. Nous supposons le métal assez bien décrit par le modèle de Sommerfeld du gas d'électrons libres. Dans l'approximation des fonction d'ondes à un électron, justifiée par la théorie de BOHM et PINES,[8] nous intégrons l'équation de Schrödinger pour ce potentiel et nous en déduisons le rapport ρ/ρ_0 de la densité électronique ρ au positon à la densité ρ_0 dans le métal non perturbé ($\rho_0 = 3/4\pi a^3$, si a désigne le rayon de la sphère électronique dans le métal). En replaçant dans l'équation (1) ZN par la densité électronique réelle au positon, on obtient le taux d'annihilation :

$$\lambda \text{ sec}^{-1} = \rho\pi r_0^2 c = \rho_0(\rho/\rho_0)\pi r_0^2 c = \frac{1{,}2 \cdot 10^{10}}{a^3} \cdot \frac{\rho}{\rho_0},$$

quand on exprime a en unités atomiques. Le temps de vie correspondant sera :

$$\tau = \lambda^{-1} = \frac{a^3}{1{,}2(\rho/\rho_0)} \cdot 10^{-10} \text{ sec.} \qquad (2)$$

Nous utilisons dans cet article les unités atomiques $e = m = \hbar = 1$.

Le potentiel d'interaction qui nous a paru correspondre le mieux à la réalité est un potentiel coulombien avec écran, de la forme $-(1/r) \cdot e^{-qr}$. Malheureusement, il exige un important travail d'intégration numérique dans chaque cas particulier. Nous n'avons effectué les calculs jusqu'au bout que pour le cuivre. Le résultat que nous obtenons ($0{,}7 \cdot 10^{-10}$ sec) paraît assez bon, comparé aux valeurs expérimentales, peu concordantes entre elles, données par BELL et GRAHAM ($1{,}2 \pm 0{,}5 \cdot 10^{-10}$ sec) d'une part[2] et GERHOLM ($2{,}7 \pm 0{,}3 \cdot 10^{-10}$) sec) d'autre part.[4]

Dans le but de montrer que le phénomène invoqué (polarisation du gaz d'électrons) suffit à donner l'ordre de grandeur des temps de vie observées, nous considérons en fin d'article le cas d'un puits carré de potentiel, pour lequel les calculs se font exactement de façon simple.

4. CAS D'UN POTENTIEL COULOMBIEN AVEC ÉCRAN AUTOUR DU POSITON DANS LE CUIVRE

Nous supposons qu'un électron à la distance r du positon voit un potentiel de la forme :

$$V = -\frac{1}{r} \cdot e^{-qr}.$$

Nous sommes alors conduits (voir en annexe) à résoudre une équation d'onde à un électron avec la masse réduite $\frac{1}{2}$. Dans le développement des fonctions d'ondes électroniques en ondes sphériques autour du positon, nous n'avons à considérer que les états s, puisque ce sont les seuls à avoir une densité non nulle à l'origine. Pour ces états, en désignant par Φ le produit par r de la fonction d'onde d'espace Ψ, l'équation se réduit à :

$$\frac{d^2\Phi}{dr^2}+\left(k^2+\frac{1}{r}\cdot e^{-qr}\right)\Phi = 0. \qquad (3)$$

Comme dans les problèmes de diffusion, nous considérons que la normalisation des fonctions d'onde à grande distance est la même que s'il n'y avait pas de perturbation; le nombre total des états devant être le même que dans le métal non perturbé et la densité des états s par unité de k étant constante, on voit que k doit varier dans l'équation (3) précisément jusqu'à la valeur k_F correspondant à l'énergie de Fermi du métal : $E_F = {k_F}^2/2$.

En divisant par deux tous les termes de l'équation (3), on constate que la *condition d'écran à réaliser est la même que pour attirer une unité de charge autour d'un centre d'impureté fixe de charge* $\frac{1}{2}$. On déduit des courbes construites à ce sujet par FRIEDEL[9] la valeur du paramètre d'écran q. Avec ${k_F}^2/2 = 0{,}26$, on trouve ainsi pour le cuivre $q = 0{,}7$.

Les solutions de l'équation (3) à retenir doivent être nulles à l'origine. Aux grandes distances, où l'influence du potentiel perturbateur est négligeable, on retrouve, à un déphasage près, les solutions en sin kr du gaz d'électrons non perturbé. Par suite le rapport de la densité électronique au positon pour un état de k donné à la densité correspondante dans le gaz non perturbé s'obtient simplement en prenant le carré du rapport des pentes de la courbe $\Phi_k(r)$ à l'origine et en un nœud r_0 assez éloigné. En effet :

$$\frac{\Psi^*\Psi_{(r=0)}}{\Psi_0^*\Psi_{0(r=0)}} = \frac{\Phi'^2_{(r=0)}}{(\text{const.}\times k)^2} = \frac{\Phi'^2_{(r=0)}}{\Phi'^2_{(r=r_0)}}$$

si

$$\Phi(r_0) = 0 \text{ et } \Phi(r) \simeq \text{const.}\times\sin(kr+\eta_0)$$

$$\text{pour } r \sim r_0.$$

On constate, en intégrant l'équation (3), que l'influence du potentiel ne se fait pratiquement plus sentir au-delà du premier nœud de Φ autre que l'origine (c'est-à-dire après quelques distances interatomiques), pour les valeurs de k qui ne sontpas extrêmement petites. Quand $k = 0$, la solution pour Φ tend asymptotiquement vers une droite à pente positive et ne présente pas de nœud, indiquant donc qu'il n'y a *pas* d'état lié. Le tableau 1 reproduit les valeurs de $\Psi^*\Psi/\Psi_0^*\Psi_0$ obtenues pour quelques valeurs de k^2.

En sommant sur tous les états, on trouve un rapport de densités ρ/ρ_0 d'environ 23; porté dans l'équation (2), ceci donne un temps de vie voisin de $0{,}7 \cdot 10^{-10}$ sec, un peu inférieur aux valeurs expérimentales. Le tableau 1 montre que les électrons très lents ($k_F^2 \leqslant 0{,}1$) sont fortement concentrés sur le positon; ils fournissent ainsi une contribution importante au taux d'annihilation. Certains facteurs négligés dans notre calcul (chocs subis par le positon et son énergie de zéro dans le potentiel du réseau) doivent élargir les niveaux électroniques, donc diminuer un peu cette contribution. Ces corrections affectent l'état initial de l'annihilation; d'autres facteurs d'élargissement, un effet Auger sur le trou positif produit dans le gaz de Fermi par l'annihilation par exemple, affectent seulement l'état final. Ils ne modifient donc pas le taux d'annihilation total, qui ne dépend que de l'état initial. Ils ont par contre un effet sensible sur la répartition angulaire des photons γ, que nous étudions au paragraphe suivant.

Enfin on attibue en général aux électrons d, dont nous n'avons pas tenu compte, les annihilations avec moment résultant supérieur à k_F. On constate expérimentalement que leur contribution est assez faible. Ceci s'explique de fait que leur densité n'est pas très forte dans les régions loin des noyeaux où se trouve le positon; et parce qu'étant fortement liés ils sont moins facilement polarisés que les électrons de valence. Leur influence ne doit donc diminuer que peu le temps de vie que nous avons calculé.

Tableau 1. *Rapport de la densité électronique au positon à la densité dans le métal pur pour diverses valeurs de* k^2

k^2	0	0,01	0,04	0,09	0,17	0,34	$0{,}52 = k_F^2$
$\Psi^*\Psi(r=0)$ / $\Psi_0^*\Psi_0$	130	52	21	12	8	5,6	4,3

5. EFFET AUGER ET CORRELATIONS ANGULAIRES DES PHOTONS

L'annihilation du positon s'accompagne de l'apparition d'un trou positif dans la bande de valence du métal. L'état correspondant ayant un temps de vie fini, il en résulte un élargissement des niveaux d'énergie par l'effet Auger. En utilisant la formule que LANDSBERG[10] a établie pour le calcul des bandes d'émission X du sodium, on obtient pour demi-largeur d'un niveau de k donné :

$$\Delta(k)=0{,}16\left\{\left(\frac{k}{\lambda}-\frac{3\lambda}{4k}\right)\text{arc tg}\left(\frac{2k}{\lambda}\cdot\frac{k_F^2-k^2}{k_F^2+3k^2+\lambda^2}\right)+\right.$$
$$+\frac{k_F^2-k^2}{2\lambda^2}\left[\frac{k_F^2-k^2}{2k\lambda}\text{arc tg}\frac{2k\lambda}{k_F^2-k^2+\lambda^2}-1\right]+$$
$$\left.+\log\frac{(k_F^2-k^2)^2+2\lambda^2(k_F^2+k^2)+\lambda^4}{4\lambda^2k^2+\lambda^4}\right\}.$$

λ est un paramètre d'écran pour les interactions électron–électron, supposées de la forme $1/r\,.\,e^{-\lambda r}$. Pour le déplacement d'une charge unité, les valeurs des paramètres d' écran sont sensiblement les mêmes, qu'il s'agisse d'une attraction ou d'une répulsion.[9] Nous avons donc pris pour le calcul $\lambda = q = 0{,}7$ pour le cuivre. On sait que la probabilité par unité d'énergie pour que le niveau E soit occupé par suite de l'élargissement $\Delta(E_1)$ d'un niveau E_1 est :

$$\frac{1}{2\pi}\cdot\frac{\Delta(E_1)}{(E-E_1)^2+[\Delta(E_1)/2]^2}.$$

Si $C = k\,n_0(E)$ désigne la densité constante des états s par unité de k dans le métal non perturbé, on aura pour densité des états par unité d'énergie $E = k^2/2$ en tenant compte de l'élargissement :

$$n(E)=\frac{c}{2\pi}\int_0^{k_F}\frac{\Delta(k_1)\,dk_1}{[(k^2-k_1^2)/2]^2+[\Delta(k_1)/2]^2}. \quad (4)$$

Finalement, le rapport ρ/ρ_0 de la densité électronique totale au positon à la densité moyenne dans le métal doit être :

$$\frac{\rho}{\rho_0}=\frac{\int_0^{E_F}\Psi^*\Psi(r=0)n(E)\,dE}{\int_0^{E_F}\Psi_0^*\Psi_0 n_0(E)\,dE}$$
$$=\frac{1}{Ck_F}\int_0^{k_F}\frac{\Psi^*\Psi(r=0)}{\Psi_0^*\Psi_0}nk\,dk. \quad (5)$$

Ayant calculé $n(E)$ par l'équation (4) pour différentes valeurs de k, l'intégration de l'équation (5) se fait numériquement, connaissant les valeurs correspondantes de $\Psi^*\Psi_{(r=0)}/\Psi_0^*\Psi_0$ obtenues par intégration de l'équation (3). On obtient ainsi pour le cuivre $\rho/\rho_0 \simeq 8{,}3$.

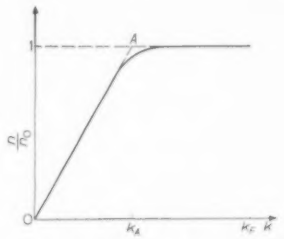


FIG. 1. Modification par effet Auger de la densité des niveaux par unité de k.

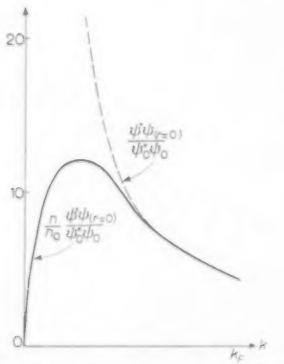


FIG. 2. Influence de la correction pour effet Auger sur la densité électronique au positon des états d'énergie positive. En pointillé: densité sans effet Auger. En trait continu: avec effet Auger. Ces densités sont exprimées en valeurs relatives par rapport aux valeurs dans le gaz non perturbé.

Les figures 1 et 2 illustrent le caractère de la correction, qui est de réduire considérablement la densité à l'origine des états de faible énergie.

L'abscisse k_A du point A (Fig. 1) fixe l'importance de la correction Auger. Un calcul simple montre que :

$$k_A = (2\Delta_0)^{\frac{1}{2}},$$

où Δ_0 est l'élargissement Auger au bas de bande, soit :

$$\Delta_0 = 0{,}32\left[\log\left(\frac{k_F^2}{\lambda^2}+1\right)-\frac{k_F^2}{k_F^2+\lambda^2}\right].$$

On voit que k_A est lié au paramètre d'écran λ de l'effet Auger.

La courbe de la figure 2 donne la densité par unité de moment calculée pour les électrons participant à l'annihilation. On voit que la densité n'est *pas du tout constante* pour les faibles moments. Il faudrait en principe la comparer aux courbes expérimentales, telles qu'on les déduit des corrélations angulaires des deux photons γ émis.[6]

Il convient de noter que dans l'équation (5) nous ne tenons pas compte des états d'énergie négative donnés par l'équation (4). Le taux d'annihilation intégré sur tous les électrons n'est pas affecté par l'effet Auger. La densité électronique de ces états au positon doit donc être égale à la diminution de densité électronique des états d'énergie positive Quand on passe, figure 2, de la courbe en trait pointillé à la courbe en trait plein, ces états liés, de moment nul, doivent donner un pic aigu à l'origine (non représenté sur la figure 2) dans la distribution angulaire des photons γ d'annihilation. Le pouvoir de résolution limité des dispositifs expérimentaux ne permet malheureusement pas de vérifier ce point.

6. APPROXIMATION DU PUITS CARRÉ DE POTENTIEL

Nous prenons pour rayon du puits celui d'une sphère électronique et nous déterminons sa profondeur par la condition que la charge électronique totale q attirée soit d'une unité. Cette charge s'exprime, en fonction des déphasages introduits au niveau de Fermi par le potentiel perturbateur, par la relation :[11]

$$q = \frac{2}{\pi}\sum_{l=0}(2l+1)\eta_l.$$

Connaissant ces déphasages en fonction de la profondeur du puits,[11] on détermine la profondeur $k_0^2/2$ convenable pour que la charge déplacée q soit égale à 1. Pour $k_F a = \sqrt[3]{(g\pi/4)} = 1{,}92$ (valeur commune à tous les métaux dans le modèle de Sommerfeld), on déduit des courbes de de FAGET et FRIEDEL[12]

$$k_0 a = 1{,}89.$$

Cette valeur de $k_0 a$ étant supérieure à $\pi/2$, il lui correspond un "état lié" auquel la grossièreté du modèle ne permet d'ailleurs pas d'attribuer une signification physique particulière.† Si Ψ est la fonction d'onde de cet état lié et si on pose $r\Psi = \Phi$, on a :

$$\Phi = \frac{1}{\sqrt{(2\pi a)}} \cdot \frac{\sin \mu r}{\sqrt{[1+(1/\lambda a)]}} \qquad \text{pour } r < a,$$

$$\Phi = \frac{1}{\sqrt{(2\pi a)}} \cdot \frac{\sin \mu a e^{\lambda a} e^{-\lambda r}}{\sqrt{[1+(1/\lambda a)]}} \qquad \text{pour } r < a,$$

avec :

$$\mu a = k_0 a \sin \mu a = 1{,}825,$$

et :

$$\lambda a = -k_0 a \cos \mu a = 0{,}48.$$

Donc, la densité à l'origine de la charge due à l'état lié est inversement proportionnelle au rayon a de la sphère électronique du métal considéré. Pour le cuivre : $a = 2{,}66$; $k_F = 0{,}725$; $\Phi_{r<a} = 0{,}14 \sin \mu r$ et $\Psi^*\Psi_{(r=0)} = 0{,}065$, soit, en comptant deux électrons par état, dix fois la densité électronique dans le métal pur.

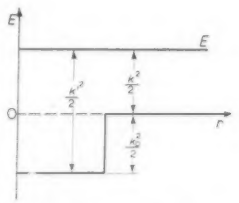


FIG. 3. Schéma du puits carré de potentiel.

En ce qui concerne les états non liés, à la fonction d'onde sin kr de l'état s correspondant à une valeur donnée de k se substituent $A \sin k'r$ pour $r < a$ et sin $(kr+\eta_0)$ pour $r > a$, le déphasage η_0 étant donné par :

$$k \operatorname{cotg}(ka+\eta_0) = k' \operatorname{cotg} k'a,$$

$$\text{avec } k'^2 = k^2+k_0^2 \text{ (Fig. 3).}$$

La densité des états s par unité de k étant constante, le rapport de la densité électronique au

† Un calcul plus exact devrait tenir compte de ce que le potentiel agissant sur l'état lié est différent de celui agissant sur les électrons de conductibilité.

positon (due aux états non liés) à la densité dans le gaz non perturbé est :

$$\frac{\rho_1}{\rho_0} = \frac{\int_0^{k_F} (Ak'/k)^2\, dk}{\int_0^{k_F} dk} = \frac{\int_0^{k_F a} (Ak'a/ka)^2\, d(ka)}{\int_0^{k_F a} d(ka)}$$

la même pour les divers métaux décrits par le modèle de Sommerfeld. La figure 4 montre l'accumulation de charge suivant les valeurs de k

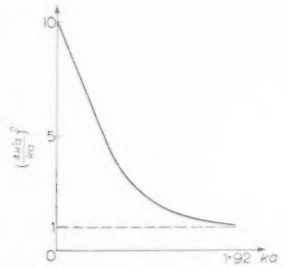


FIG. 4. Densité électronique due aux états non liés dans l'approximation du puits carré de potentiel.

jusqu'à l'énergie de Fermi du métal. Par intégration numérique, on trouve $\rho_1/\rho_0 = 3{,}7$. Ce schèma du puits carré de potentiel conduit donc au total à une densité électronique au positon quatorze fois plus forte environ, pour le cuivre, que dans le métal non perturbé. Ceci donne un temps de vie de $1{,}1 \cdot 10^{-10}$ sec en bon accord avec la valeur expérimentale $1{,}2 \cdot 10^{-10}$ sec donnée par BELL et GRAHAM[2] mais pas avec celle, plus récente, de $2{,}7 \cdot 10^{-10}$ sec obtenue par GERHOLM.[4]

Plus généralement, en additionnant la densité de l'état "lié" (inversement proportionnelle à a) et celle des états non liés, on obtient pour le taux d'annihilation λ et le temps de vie τ, respectivement, les formules suivantes :

$$\lambda = 1{,}2\left[\frac{1{,}45}{a} + \frac{3{,}7}{a^3}\right] 10^{10}\ \text{sec}^{-1};$$

$$\tau = 1/\lambda = \frac{10^{-10}}{1{,}2} \cdot \frac{a}{1{,}45 + (3{,}7/a^2)}\ \text{sec.} \qquad (6)$$

Dans le tableau 3, nous comparons les temps de vie déduits de ces formules aux valeurs expérimentales connues pour différents métaux.

On constate que ce modèle *extrêmement simplifié* de puits de potentiel fournit l'ordre de grandeur correct en faisant appel uniquement à la polarisation du gas électronique dans le champ du positon. Son insuffisance se manifeste cependant en particulier en ce que les formules (6) font trop dépendre le temps de vie de la nature du métal (par a pris assez arbitrairement égal au rayon de la sphère électronique). Cet effet se fait sentir principalement pour l'aluminium, à cause de ses trois électrons de valence, et dans la progression qu'on obtient avec les alcalins alors que les valeurs expérimentales sont nettement plus constantes. Nous tenons à remercier les Drs. W. KOHN et R. A. FERRELL pour des critiques constructives, en particulier sur le rôle de l'effet Auger.

ANNEXE

Choix des fonctions d'ondes à un électron dans l'interaction coulombienne avec écran

Soient (x_0, y_0, z_0) les coordonnées du positon dans un système d'axes orthogonaux fixes dont nous supposons l'origine au centre de gravité du gaz d'électrons; soient (x_i, y_i, z_i) les coordonnées du $i$ème électron dans le même système. Puisque nous considérons le cas d'interactions centrales électrons-positon, posons :

$$\xi_i = x_i - x_0; \qquad \eta_i = y_i - y_0; \qquad \zeta_i = z_i - z_0.$$

Tableau 2. Comparaison des temps de vie observés et calculés par l'équation (6)

Métal	$\tau_{obs} \cdot 10^{10}$sec	$\tau_{calc} \cdot 10^{10}$sec	Métal	$\tau_{obs} \cdot 10^{+10}$sec	$\tau_{calc} \cdot 10^{+10}$sec
Li	$1{,}5 \pm 0{,}6$[2]	1,5	Cu	$1{,}2 \pm 0{,}5$[2]	1,1
Na	$1{,}5 \pm 0{,}6$[2]	2	Cu	$2{,}7 \pm 0{,}3$[4]	
K	$1{,}5 \pm 0{,}6$[2]	2,5	Al	$1{,}5 \pm 0{,}3$[2]	0,8
Ag	$1{,}5 \pm 0{,}6$[2]	1,3	Al	$2{,}5 \pm 0{,}3$[4]	
Cs	$1{,}5 \pm 0{,}6$[2]	2,9	Au	$1{,}2 \pm 0{,}3$[2]	1,3
			Hg	$1{,}2 \pm 0{,}5$[2]	1,5

VOL. 4 1958

Le hamiltonien :

$$\mathscr{H} = -\tfrac{1}{2}\left[\Delta_{(x_0,y_0,z_0)} + \sum_{i=1}^{N} \Delta_{(x_i,y_i,z_i)}\right] + \sum_{i=1}^{N} V_i,$$

devient, avec les nouvelles variables :

$$\mathscr{H} = -\sum_{i=1}^{N} \left\{\Delta_{(\xi_i,\eta_i,\zeta_i)} + \right.$$

$$\left. + \sum_{j>i}^{N}\left[\frac{\partial^2}{\partial\xi_i\partial\xi_j} + \frac{\partial^2}{\partial\eta_i\partial\eta_j} + \frac{\partial^2}{\partial\zeta_i\partial\zeta_j}\right] - V_i\right\}.$$

Par suite de l'apparition des termes croisés, les variables ne se séparent pas. Cependant, la théorie de BOHM et PINES[8] fait prévoir qu'on obtiendra une bonne approximation au moyen de fonctions d'un électron, pourvu qu'on ait déterminé de façon convenable le paramètre d'écran q dans $V_i = (1/r_i)e^{-qr_i}$. Les solutions Ψ_i qui ne dépendent que de la position relative du $i$éme électron par rapport au positon seront données par :

$$[-\Delta_{(\xi_i,\eta_i,\zeta_i)} + V_i]\Psi_i = E_i\Psi_i, \qquad (3')$$

C'est l'équation de Schrödinger du système électron-positon avec la *masse réduite* $\frac{1}{2}$. Notons que l'énergie du système pris dans son ensemble ne sera pas la somme des E_i à cause des termes croisés dont les contributions ne sont pas nulles pour la fonction d'onde totale.

REFERENCES

1. HEITLER W. *The Quantum Theory of Radiation*. University Press, Oxford (1954).
2. BELL R. E. and GRAHAM R. L. *Phys. Rev.* **90,** 644 (1953).
3. DE BENEDETTI S. and RICHING H. J. *Phys. Rev.* **85,** 377 (1952).
4. GERHOLM T. R. *Ark. Fys.* **10,** 523 (1956).
5. MINTON G. E. *Phys. Rev.* **94,** 758 (1954).
6. FERRELL R. A. *Rev. Mod. Phys.* **28,** 308 (1956).
7. LEE WHITING G. E. *Phys. Rev.* **97,** 1557 (1955).
8. PINES D. *Solid State Physics* (Ed. F. SEITZ and D. TURNBULL) Vol. 1, p. 367. Academic Press, New York (1955).
9. FRIEDEL J. *Advances in Phys.* **3,** 446 (1954).
10. LANDSBERG P. T. *Proc. Phys. Soc.* A **62,** 806 (1949).
11. FRIEDEL J. *Phil. Mag.* **43,** 153 (1952).
12. DE FAGET DE CASTELJAU P. et FRIEDEL J. *J. Phys. Radium* **17,** 27 (1956).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 118–127.

HALL EFFECT AND ELECTRICAL CONDUCTIVITY OF TRANSITION-METAL DIBORIDES*

H. J. JURETSCHKE* and R. STEINITZ†

* Polytechnic Institute of Brooklyn, Brooklyn, New York

† American Electro-Metal Division of Firth Stirling Inc., Yonkers, New York

(*Received* 17 *June*, 1957)

VOL. 4 1958

Abstract—The room-temperature resistivity and Hall constant of the diborides of the fourth- to sixth-column transition metals and their solid solutions have been measured in order to clarify the electronic structure of these compounds. Since the experimental samples, prepared from powders, were usually porous, extrapolation procedures have been developed to relate the electrical properties measured on porous samples to those to be expected at full density. The results thus extrapolated indicate that all these compounds have a similar band structure, the substitution of one transition metal for another changing mainly the occupation of the conduction band. This conduction band is of standard shape; it is nearly empty for fourth-column transition metals, and is filled to about one electron per cell upon substituting a fifth-column metal. If the band structure of these compounds derives from that of the transition metals, the results suggest that the usual *s–d* band overlap is here considerably reduced, and that the *d*-bands are completely filled by the contribution of three electrons from each boron atom. This makes the main conduction band an *s*-band. Such an interpretation is in accord with a number of other characteristic properties of these compounds.

1. INTRODUCTION

THE refractory borides of the transition metals of the fourth to sixth groups of the Periodic System have attained considerable importance during the last years as constituents of various high-temperature materials. Many of their physical and mechanical properties have been investigated in detail and are rather well known today.[1] However, though the crystal structure of these substances is simple, the electron distribution between metal and non-metal atoms and the kind of bonding between them is still obscure. The present work was undertaken to clarify the electronic structure of the diborides by measurements of their electrical resistivity and Hall effect at room temperature.†

The diborides of the fourth-, fifth-, and sixth-group transition metals form isomorphic compounds in the simple hexagonal structure shown in Fig. 1. They have very high melting points, mostly higher than that of the parent metal, they are chemically very stable, and they show surprisingly metallic electrical properties. Knowledge of their electronic structure would provide the basis for an understanding of these characteristic properties of the compounds. Furthermore, experimental information on the electronic structure

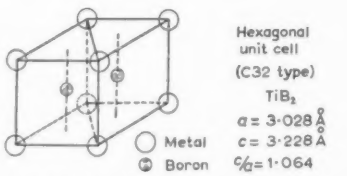


FIG. 1. Crystal structure of the transition - metal diborides. The lattice parameters refer to TiB_2. A c/a value near unity is typical of all these compounds.

* This work has been supported in part by the U.S. Office of Naval Research.

† A preliminary account of this work was presented at the January 1956 meeting of the American Physical Society (see *Bull. Amer. Phys. Soc.* **1**, 54 (1956)).

would bear on the long-standing question whether in metal–non-metal compounds of this kind the non-metal acts as an acceptor or donor of conduction electrons. A good summary of the arguments and evidence for both alternatives has been given by SCHWARZKOPF and KIEFFER.[1] Although, in fact, both points of view can account for many properties of the diborides, the underlying models are quite distinct. Thus, the crystal structure is thought of either as a typical three-dimensional boron net, with strongly directional bonds, strengthened by the metal atoms; or as a simple hexagonal metallic crystal stabilized by the interstitial boron atoms, the electrons of the boron being shared to a considerable extent by the metal.

Measurements of the electrical properties, and in particular of the Hall effect, can be expected to shed some light on the bonding role of the two constituents in these compounds. From the results for any one compound alone, we obtain the sign of the predominant carrier and qualitative information about its mobility and density. Furthermore, the existence of a large number of isomorphic compounds, and of their continuous solid solutions, permits a detailed study of the change of these properties as one transition metal is substituted for another. All the compounds having the same crystal structure, such substitution alters, in the first approximation, only the number of valence electrons that the metal contributes to the electronic structure of the solid. Under favourable conditions, such comparative study of the change in occupation of the conduction bands can lead to a rather detailed specification of the structure of these bands. Clearly, it also allows separating to a considerable extent the contributions of the metal and the non-metal to the conduction process.

Since the diborides are formed by a solid-state reaction at very high temperatures, the resulting homogeneous material is available only in powder form, and samples suitable for measurement are prepared by pressing the powder. Such pressed samples are porous, and since the measured electrical properties are a function of porosity, the raw experimental data of the various samples have to be extrapolated to values applying at the theoretical full density. Part of our work has been concerned with establishing extrapolation procedures for data on the resistivity and the Hall effect obtained on samples of various porosities.

This method of preparation has a basic drawback in limiting the interpretation of the experimental results, because the hexagonal anisotropy of single-crystal material is suppressed. In view of the good electrical properties of these materials, however, it is unlikely that the anisotropy is large, or that it is important in determining the electronic structure.

2. SPECIMEN PREPARATION AND EXPERIMENTAL METHOD

Most of the boride powder lots from which Hall-effect samples were prepared, and in particular TiB_2 and VB_2, were made by carbon reduction of appropriate mixtures of the metal oxides and B_2O_3. These borides were prepared by induction heating for several hours at temperatures in the neighborhood of 2000°C. Where powders could be made by the direct-synthesis method or by reaction of the metal hydride with boron, such as in the cases of CrB_2, TaB_2, and ZrB_2, the borides were formed by heating for 1 hr at approximately 1800°C. All boride powders were crushed in a stainless-steel mortar and pestle to under 100 mesh and subsequently either micronized or ball-milled for 24 hr to reduce the particle size. They were then checked for purity and composition by chemical and X-ray analysis.

In the case of solid solutions, the individual boride powders were mixed in the desired proportions, tumbled, and placed into graphite crucibles. They were heated in a high-frequency furnace for 1 hr at 1800°C; for solid solutions containing ZrB_2 and/or NbB_2, this temperature was 2250°C. Each product was then crushed to below 100 mesh and analyzed again by chemical and X-ray analysis for purity and completeness of reaction.

The X-ray analysis also provided values of the theoretical density of the various materials, necessary for applying the porosity corrections for the electrical properties. The lattice constants of the pure compounds obtained are in good agreement with the values given in the literature.[1,2] The parameters for the solid solutions are also in line with those of the constituent borides. The deviation of the lattice constants from Vegard's law is usually quite small.

For Hall-effect measurements, it was necessary to fabricate from the powder specimens in the form of long and thin rectangular bars. Since many of these borides have melting points in excess of 2500°C, the possibility of making samples by cold pressing and sintering was excluded. It was therefore necessary to hot press all samples in a graphite die directly to the final size, $5 \times 1{\cdot}2 \times 0{\cdot}5$ cm. Thinner samples having sufficient mechanical strength could not be prepared by pressing.

Hall voltages were measured with a three-probe method, using direct current, in fields up to 15 kG. Since these materials cannot be soldered or welded, all electrical contact was made by mechanical pressure alone. This limited the sample current to about 15 A before appreciable contact heating set in. Consequently, the large cross-section of the samples led to current

densities below 25 A/cm², producing very small Hall voltages in most of the compounds. Since the reproducibility of these small signals was at most good to within 10 per cent, none of the usual refinements for obtaining accurate Hall constants, such as field reversal and corrections for secondary effects, were observed. Most of the uncertainty in the measurements can be traced to the difficulty in making reliable and reproducible contact to rather porous samples. An additional source of instability was provided by the thermal drifts caused by the non-uniform heating of the sample because of local fluctuations in resistivity.

3. POROSITY-DEPENDENCE OF ELECTRICAL PROPERTIES

The solids produced by pressing the powders usually do not reach the density of single crystals; the porosity may be as high as 45 per cent. Since the electrical properties of interest here are porosity-dependent, it is necessary to know how to relate the experimental results obtained on a sample of given density to those proper at the theoretical density.

For this purpose an extensive series of measurements was carried out on TiB_2 samples covering a large range of densities. The results have led to an empirical extrapolation law for the resistivity, and have shown good agreement with the extrapolation law for the Hall effect predicted by simple theory. In order to test the general applicability of these relations to the other diborides, a similar series of measurements was carried out on VB_2 samples. The same laws were found to apply.

The experimental relationship between measured Hall constant R_m and relative density $(1-P)$ is shown for TiB_2 in Fig. 2. The most striking feature is, of course, the considerable dependence of the measured Hall constant on density.

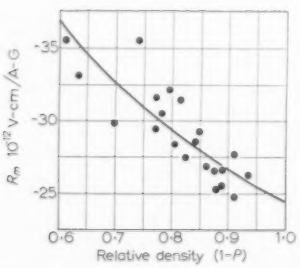


FIG. 2. Measured Hall constant R_m of TiB_2 as a function of porosity P. The full curve gives the variation predicted by equation (2).

The relation of measured Hall effect to porosity of a sample has until recently never been discussed fully in the literature. An estimate by VOLGER[3] of the Hall effect to be expected in a two-dimensional uniform mixture of two materials of differing Hall constants R_0 and resistivities ρ_0, with an appreciable difference in the two resistivities, indicates that in such a mixture the measured Hall constant would correspond to that of the pure highly conductive component. It is obvious that such an interpretation does not apply here. LANDAUER[4] has suggested that the surface charge of the pores required to guide the current around the pores is modified in the presence of the magnetic field in such a way that the pore interiors acquire an electric field component transverse to the average current flow direction and the magnetic field. This additional field would therefore be recorded as a Hall voltage. A theoretical estimate of this effect has been carried out recently[5] for porous materials in which the pores are sufficiently far apart so as to act independently. It is shown that the influence of the pores depends on their shape with respect to the directions of the electric current and magnetic field. Thus, cylindrical pores parallel to the magnetic field cause no change (in agreement with VOLGER's result), while cylinders normal to the magnetic field yield a measured Hall constant R_m:

$$R_m = \frac{R_0}{1-P}. \tag{1}$$

For spherical pores, the Hall constant changes with porosity according to the relation:

$$R_m = \frac{1-\frac{1}{4}P}{1-P} R_0. \tag{2}$$

In a porous sample pressed from powder particles of fairly uniform dimensions, one expects that pore shapes vary between spherical and elliptical, very long needles being rather unlikely. In addition, pore sizes are rather uniform, being determined mainly by the dimensions of the starting powder. These properties are confirmed by photomicrographs of samples of TiB_2 pressed from the same powder to different final densities. The main difference is the density of the pores, not their size.

Under these conditions, the measured variation

of the Hall constant with porosity P should correspond to some average of the above two results, probably weighted in favor of the formula (2) for the spheres. Actually, the theory is not very sensitive to pore shape, as long as the pores are oriented at random. For instance, the net effect for cylindrical pores oriented in equal numbers along the three principal directions is practically indistinguishable from equation (2), for $P < 0{\cdot}4$.

The best fit of equation (2) to the experimental data for TiB_2 is shown by the solid line in Fig. 2. It is apparent that the formula (2) agrees with the experimental results over a considerable range of densities. The scatter of the individual points about this curve is principally due to density gradients in the sample and more microscopic fluctuations in porosity caused by slight variations in the technique of preparation, such as the pressing temperature and the rate of compression. The possible anisotropy of the Hall constant may also play a role. The data include samples from three powder lots of somewhat different chemical composition, prepared by varying raw materials and mechanical treatment. The fluctuations due to such deliberate variations fall within those found in the group of samples prepared under, to a first approximation, identical conditions.

The resistivities of the same group of samples are shown in Fig. 3. As expected, there is a large variation of resistivity with density. In addition, the resistivity seems to be much more sensitive than the Hall constant to the method of preparation. Thus, one powder lot (No. 930, open circles) gives uniformly lower resistivities than the other lots used.

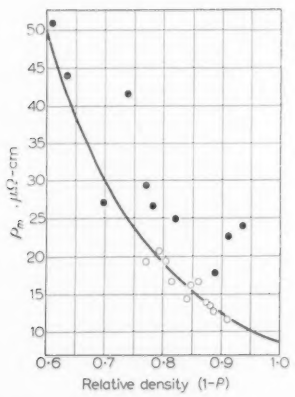


FIG. 3. Density-dependence of the resistivity ρ_m of TiB_2. All open circles refer to samples prepared from a single powder lot; the full circles represent various other lots. The line is drawn according to the empirical expression (3).

In view of the large scatter of points in Fig. 3, no functional relationship of the porosity-dependence of resistivity, based on all points in Fig. 3, can be established, However, the samples prepared from lot No. 930 show a reasonable regularity. In the absence of a good theory to explain this large porosity-dependence, we have tried to fit the data to a law of the form

$$\rho_m = (1-P)^n \rho_0.$$

For the resistivities of lot No. 930 alone, we find the best values $n = -3{\cdot}4$, $\rho_0 = 8{\cdot}7$ $\mu\Omega$-cm.

Obviously, these points do not, in themselves, constitute a very sensitive test of the best exponent, since the range of densities involved is rather small. Nevertheless, if we assume for simplicity an exponent $n = -3{\cdot}5$, and the same ρ_0, we obtain the full curve shown in Fig. 4. It is apparent that the over-all fit of most points to this curve is not unreasonable.

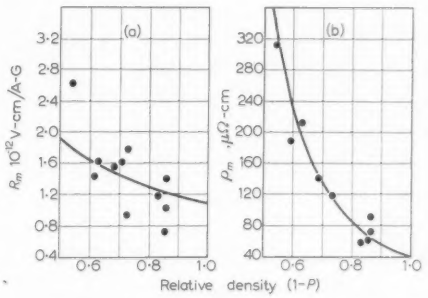


FIG. 4. Density-dependence of (*a*) the Hall constant and (*b*) the resistivity of VB_2. The full curves represent equations (2) and (3), with $R_0 = 1{\cdot}1 \times 10^{-12}$ v-cm/AG, $\rho_0 = 39 \times 10^{-6}$ Ω-cm.

These results, therefore, suggest a law of the form

$$\rho_m = (1-P)^{-3.5} \rho_0 \qquad (3)$$

for extrapolating measured resistivities to those expected at full densities. Clearly, the extrapolated values obtained from a small number of samples will show appreciable spread, especially if the

samples are prepared from different starting powders. Nevertheless, equation (3) should give some measure of the properties of fully dense material.

As a test of the applicability to other diborides of the extrapolation formulas for the measured Hall effect and resistivity established above for TiB_2, a similar set of measurements of R_m and ρ_m as a function of $(1-P)$ was carried out on a series of samples of VB_2. This diboride has a Hall constant of much smaller magnitude, of the order of the Hall constants for the better pure metallic conductors. The resistivity, on the other hand, is appreciably higher than that of TiB_2. The direction of both these changes made the measurement of the Hall effect much more difficult. Correspondingly, the accuracy of measurements involving much smaller signals (about 0·1—0·2 μV at 15kG and 12 A primary current) decreased sharply.

The results of the Hall-effect measurements on VB_2 are shown in Fig. 4(a). The points include measurements on samples from various powder lots, the accuracy of any single point lying between 10 and 20 per cent. The figure also shows the curve which represents the average description of the points by equation (2). The agreement is as good as could be expected from the data.

The resistivities of the samples of the same group are shown in Fig. 4(b). Again, points are included from various powder lots. The curve shown represents a fit of the points to a law of the form of equation (3). Although it is unlikely that the expression (3) has any general significance, the agreement obtained in Fig. 4(b) suggests that it is a reasonable extrapolation formula for the compounds under discussion here, all prepared by very similar chemical and physical methods, and lying in the same range of porosities.

Recently, SAMSONOV and NESHPOR[6] have found that the resistivity/relative density relation for a series of similar compounds can be fitted by an exponential curve $\rho/\rho_0 = \exp(A/1-P)$ over a large range of densities. If this expression is applied to VB_2, we obtain susbtantially the same value of ρ_0 as before. Hence in our range of densities both extrapolation formulas appear equally applicable.

The interpretation of the results on VB_2, both for R_0 and ρ_0, is subject to the same limitation as that discussed in connection with TiB_2. In both cases the measurements probably represent a complicated average over the pairs of constants which describe the properties of the substance on single crystals.

4. ELECTRICAL PROPERTIES OF THE TRANSITION-METAL DIBORIDES

With the exception of tungsten, all transition metals of the fourth, fifth, and sixth columns of the Periodic Table form isomorphous diborides. Tungsten and also molybdenum form borides of somewhat more complicated structure, which are usually described by the formulas W_2B_5 and Mo_2B_5. In the case of molybdenum the two forms are connected by a polymorphic transition, the normal diboride being the stable high-temperature phase.

Measurements were carried out on all these diborides. The number of samples tested for each diboride was small, being limited, in fact, to one sample in a few cases. However, since the extrapolation procedures developed above gave good agreement for duplicate samples of the same compound, in all cases where these were available, such a survey was sufficient to establish a general picture of the electrical properties of all the diborides.

Table 1 lists the extrapolated Hall and resistivity values for all the diborides, with the metals arranged according to their positions in the Periodic Table. Values shown in parentheses refer to the alternate structure of the diboride. The two values of the extrapolated resistivity of two samples of HfB_2 are given separately.

Nearly all Hall constants are negative. All three diborides of the fourth column have very large Hall constants, and in all three rows there is the same sharp decrease in the Hall constant from column IV to columns V and VI. The relative values in a given column show less regularity. It seems significant, though, that the smallest Hall constants occur in the sixth column, and that the alternate boride structure of molybdenum has a small positive Hall constant.

The resistivities are generally lower than the values of various workers summarized by SAMSONOV.[7] In most cases these earlier values have not been related systematically to the porosity of the samples. Just as the Hall constants, the resistivities also show some pattern. For a given column the central row has the lowest resistivity, while, on

Table 1. *Extrapolated electrical properties of transition-metal diborides at room temperature. Values in parentheses ( ) refer to the crystal structure* M_2B_5.

(*a*) *Hall Effect* (Units: 10^{-12} Vcm/AG)

Ti	V	Cr
−24	−1·1	−1·1
Zr	Nb	Mo
−20	−1·0	−0·5 (+0·1)
Hf	Ta	W
−17	−2·7	(−1·6)

(*b*) *Resistivity* (Units: 10^{-6} Ω-cm)

Ti	V	Cr
9	38	56
Zr	Nb	Mo
7	12	30 (18)
Hf	Ta	W
15, 8	14	(39)

the whole, the resistivity increases from left to right. The relative resistivities of the two molybdenum structures indicate that WB_2, if it exists, is probably a bad conductor.

The general pattern of the variation of the Hall constants from the fourth to the sixth column suggests the following qualitative picture of the electronic structure. If the main conduction in all these materials is due to electrons in a single energy band, the large Hall constants in the fourth column imply that the number of conduction electrons in the fourth-column diborides is rather small. Thus, we are dealing with a nearly empty band, either separated by a gap from the next lower-lying band or overlapping this band to a small extent. In the fifth column, the Hall constants are also negative, but of normal size, suggesting that the nearly empty band existing above has now been filled, to the extent of about one conduction electron per unit cell. The additional electron presumably comes from the metal.

The very small change in the Hall constant from the fifth-to the sixth-column metals might indicate that no additional electrons are contributed to the conduction band under this substitution. Alternately, if the electron population in fact continues to change, the results imply either that the conduction band is no longer of standard shape at the Fermi level, or that we are dealing with the simultaneous filling of several overlapping bands. These measurements cannot distinguish between the various possibilities.

The resistivities can also be interpreted in terms of the above picture. Using the values of Table 1, we can obtain the effective mobilities R_0/ρ_0. For the fourth-column materials, they range from 200 to 300 cm^2/V sec., while the mobilities of the remaining compounds are generally smaller by a factor of fifty or more.

The abnormally high mobilities in the fourth column are characteristic of semi-metals having less than one-quarter conduction electrons per cell[8] the electrons in this case having such long wavelengths that they will not be scattered by all lattice vibrations. Column-five and column-six diborides are relatively poor conductors. It appears from the small linear resistivity/temperature relation, that a large fraction of the measured resistivity is residual. This means that the low mobility is most likely due to imperfections and faults in the crystalline structure. Support for this hypothesis comes from the finding that column-five and column-six diborides may exist over a range of boron concentrations on either side of the stoichiometric ratio. The variation of mobility from row to row may be due to a corresponding variation in the Debye temperature of the lattice.

The identification of the location of the band edge in the fourth-column diborides, with respect to the band structure of the parent metal, would give an important clue about the type of binding and the electronic structure of the diborides. An attractive possibility is that six electrons per cell of the boron enter the conduction bands of the metal.[9] In this scheme, one would think of the diborides as transition metals with filled *d*-bands conducting in the next higher *s*-band, and one would expect an anology of their properties to those of the series, say palladium, silver, and cadmium.

5. ELECTRICAL PROPERTIES OF SOLID SOLUTIONS OF THE DIBORIDES

The transition-metal diborides readily form solid solutions, at all concentrations, in all of the

large number of combinations tested and, in particular, for neighbors in the same row.[10]

Because of this possibility, a number of the unresolved questions concerning the electronic structure of these compounds suggested by the measurements on the pure materials can be investigated in detail. Among these are the electron-donor characteristics of the various metals, the possibility of conduction by more than one kind of carrier, and the shape of the conduction band.

It is convenient to discuss the results of the Hall-effect measurements for the solid solutions in terms of the effective density of the carriers n^*, defined by

$$n^* = \frac{1}{ecR_0}. \qquad (4)$$

The X-ray density of each solid solution needed to find n^* has been obtained either by direct measurement or by interpolation, using Vegard's law. In all cases it is assumed that the solution is complete, and that the concentration is that of the original mixture. Both assumptions have been tested and verified on a number of samples.

(*a*) $(Ti_x, V_{1-x})B_2$ *and* $(Zr_x, Nb_{1-x})B_2$

As suggested from the data on the pure diborides a solution of column-four to column-five metals should result in an orderly decrease in the Hall constant as x, the atomic concentration of titanium or zirconium, is decreased. This does, in fact, occur, making it reasonable to assume that the conduction-electron population changes appreciably as the concentration of column-five metal increases.

The extrapolated resistivity of the solid solutions $(Ti_x, V_{1-x})B_2$ and $(Zr_x, Nb_{1-x})B_2$ is shown in Fig. 5. It exhibits the normal resistivity/concentration curve of a disordered solid solution, giving a maximum additional resistivity near $x = 0{\cdot}5$.[11] It is worth noting that the maximum actually occurs somewhat to the left of this value. This is indicative of an increase in the density of the conduction electrons, with decreasing value of x. The shift of the resistivity maximum is seen to be more pronounced in the resistivity of the $(Zr_x, Nb_{1-x})B_2$ alloys, since the resistivities of the end-point compounds are somewhat more alike.

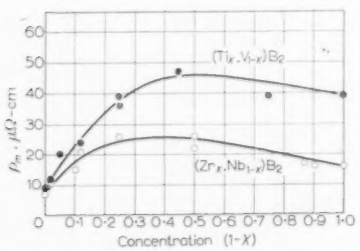


FIG. 5. Resistivity of the solid solutions $(Ti_x, V_{1-x})B_2$ and $(Zr_x, Nb_{1-x})B_2$. The data represent the extrapolation of the measured values to 100 per cent density according to equation (3).

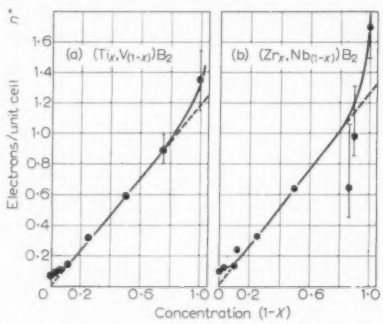


FIG. 6. Effective electron density n^* for $(Ti_x, V_{1-x})B_2$ and $(Zr_x, Nb_{1-x})B_2$.

The change in the carrier concentration of these solid solutions deduced from the Hall constants is shown in the variation of n^* with $(1-x)$ in Fig. 6. The graphs show that the interpretation of the change of Hall constant as a reflection of the change of carrier population is more than qualitative. Over a large range of values of $(1-x)$, both graphs are characterized by the following features: (a) n^* varies linearly with $(1-x)$, (b) the slope is close to unity, and (c) the extrapolation of the straight-line portion goes through the origin.

All these characteristics, taken together, are strong evidence that, for a considerable range of values of x, the conduction band is of standard shape and conduction is due to one carrier alone. Furthermore, the regular increase of n^* with addition of column-five metal, with a rate of nearly unity, means that the additional metal electrons go fully into the conduction band. The deviation from unity of the slope is probably a measure of anisotropy of the conduction process.

The scattering of the points in Fig. 6 can be

VOL. 4 1958

ascribed to the fact that different samples came from different powder lots and were prepared at different times. This is of particular importance in the zirconium–niobium series, since NbB_2 can exist over a range of concentrations of boron away from the stoichiometric ratio. An excess or deficiency of boron would have great importance if the valence electrons of boron help to determine the density of conduction electrons. The probable errors of the various points are indicated in Fig. 6. At low values of $(1-x)$, they are very small; at large values, they have to be taken into consideration.

The deviations from a straight-line relationship between n^* and x occur near $x = 0$ and near $x = 1$. That near $x = 1$ takes place in a direction and a manner which indicate that in this region conduction is due to more than one type of carrier. The Hall constant of two overlapping bands of standard shape, containing n_1 electrons and n_2 holes per unit volume, is given by

$$R = \frac{1}{ec} \cdot \frac{n_1\mu_1^2 - n_2\mu_2^2}{(n_1\mu_1 + n_2\mu_2)^2} \qquad (5)$$

where μ_1 and μ_2 are the respective mobilities. It is clear from equation (5) that the Hall constant is not going to increase as fast as $1/n_1$, as n_1 is decreased, if the second term in the numerator appears. This is in agreement with the change in n^* at low values of $(1-x)$ in Fig. 6. However, the effect is appreciable only for small values of n_1 and n_2. We can conclude from both the size of the Hall constants and the direction of deviation of n^* from the straight-line relation, that near $x = 1$ two bands are overlapping to a small extent, the higher band being of standard shape. Since at $x = 1$ we have $n_1 = n_2$ (the straight line goes through the origin), the Hall constant at $x = 1$, i.e. for pure TiB_2 or ZrB_2, must be interpreted according to equation (5):

$$R(x = 1) = \frac{1}{nec} \cdot \frac{\mu_1 - \mu_2}{\mu_1 + \mu_2}. \qquad (6)$$

From the sign of R, we know that $\mu_1 > \mu_2$, and from the deviation of n^* from the straight line we also know that μ_2 is not negligible. Hence the real degree of overlap between bands is certainly less, probably by a factor of two or more, than that indicated by the effective n^* for these compounds. This puts them into the same class as the semimetals bismuth and antimony.

The deviation of n^* from a straight line at high values of $(1-x)$ does not have as clear-cut an explanation. At large values of n^* one may expect anomalously low Hall constants either because of overlap of bands, as indicated by equation (5), or because a single band is no longer of standard shape.

Other possibilities are that the band structure itself changes appreciably as the atomic concentration of the metals is changed at high $(1-x)$, or that impurity scattering effects, such as have been discussed by COLES,[12] become important in this range. Under such circumstances the explanation of the experimental results would certainly not be simple.

(b) Other solid solutions

In order to investigate whether the method of interpretation used in the preceding section is also applicable to other solid solutions of diborides, measurements have been carried out on a limited number of concentrations of the compounds $(V_x, Cr_{1-x})B_2$, $(Ti_x, Cr_{1-x})B_2$, $(Zr_x, Ta_{1-x})B_2$, and $(Zr_x, Mo_{1-x})B_2$.

The difficulty in obtaining reliable Hall measurements for most of these samples had the same sources as those already discussed in connection with the pure compounds, except that the resistivity was in most cases even more unfavourable, owing to high residual resistivity. Furthermore, the small number of samples tested leaves an uncertainty as to whether the samples were in fact representative.

Thus, in the vanadium–chromium series, the error in the values of n^* is probably as large as 30 per cent. Within this uncertainty it appears as if in this series n^* remains constant, so that we can conclude that there is no drastic change in the electron population n^* in going from column-five to column-six metals.

A similar uncertainty attaches to the interpretation of the data of the $(Ti_x, Cr_{1-x})B_2$ series. It was found that various groups of samples, prepared at different times, gave very different results. The data presented in Fig. 7 are among the most reliable. As can be seen, the values cluster about the solid line which was drawn on the assumption that chromium contributes two electrons to each one of

vanadium in the filling of the conduction band. However, at small concentrations, where the reproducibility of the properties in different samples was best, the deviation towards a lower effective

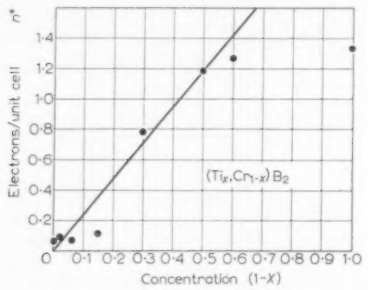


FIG. 7. Effective electron density n^* for $(Ti_x, Cr_{1-x})B_2$. The slope of the solid line corresponds to the filling in $(Ti_x, V_{1-x})B_2$, assuming that each chromium atom contributes two electrons to the conduction band for each one of vanadium.

valency of chromium is appreciable. On the other hand, the data seem to confirm the behaviour of the vanadium–chromium solid solutions by showing little variation in n^* for $(1\text{-}x) > 0{\cdot}5$.

A small number of solid solutions having ZrB_2 as a starting point have been measured in order to test whether the electron additivity also applies from one row to the next, as in $(Zr_x, Ta_{1-x})B_2$, and whether molybdenum has similar characteristics to chromium. The results show that, within the range of concentrations measured ($1{\cdot}0 > x > 0{\cdot}88$), the substitution of tantalum for niobium in $(Zr_x, Nb_{1-x})B_2$ gives precisely the same values of n^* as are shown in Fig. 6(b). Similarly, near $x = 1$, molybdenum in $(Zr_x, Mo_{1-x})B_2$ has the same effect as chromium in $(Ti_x, Cr_{1-x})B_2$ (Fig. 7). We can therefore conclude that the band structure is essentially the same from one row of transition metals to the next and changes in any one row from column four to column six in similar manner.

In an attempt to penetrate the two-band region, and thus to obtain more information about the lower-lying band, we tried to produce solid solutions of titanium diboride with aluminum diboride. Aluminum is not a transition metal, and the existence of a diboride of it has been in doubt for a long time. Recently, its existence has been confirmed unambiguously,[13] but the preparation caused considerable difficulties. Since it proved impossible to hot press AlB_2 pieces without decomposition of the boride, it was decided to try to produce solid solutions of TiB_2 and AlB_2 by the addition of aluminum and boron in the correct ratio to TiB_2 powder. This mixture was reacted at around 1400°C, and sample pieces were hot pressed from it. There was no difficulty in the production of these samples, and Hall measurements were taken on them. However, X-ray and microscopic examination revealed that no solid solutions had formed, the samples consisting of mixture of TiB_2 and an unknown compound containing aluminum.

6. DISCUSSION

The experimental results suggest a very simple band structure for these compounds. The conduction band has free-electron properties for a considerable range of filling. At low energies it overlaps slightly with another band, of somewhat higher effective mass and density of states, while above filling to about one electron per cell, it either deviates from standard shape or starts to overlap with other, higher-lying, bands. The relative electrical properties of the diborides of the various transition metals result mainly from the number of electrons the metal atom contributes to the conduction band, this number increasing in an orderly fashion from the group-four to the group-six metals.

This description ignores the anisotropy that all conduction processes may be expected to have in hexagonal crystals. However, since these compounds are good metallic conductors, the anisotropy cannot be severe. In the discussion below, we assume that it can be neglected in explaining the gross features of the band structure.

Without doubt, the strongly metallic properties of these compounds must derive from those of the parent metal. Since the interatomic spacing of the metal atoms is practically the same in the compounds as in the pure metal, the band structure of the metal atoms alone, in a simple hexagonal arrangement, will be similar to that of the parent metal. The introduction of the boron into the interstices of this simple hexagonal metallic structure can have two extreme alternate effects: either it leads to a complete reordering in energy of the original band structure, introducing new gaps, and some new, predominantly boron, bands; or else it results mainly in a different filling of the existing bands without their major features being changed. The second alternative is more probable. The metal bands are predominantly *d*-bands, covering a range of about 10 eV with a characteristic high

VOL. 4 1958

density of states. The boron atoms, having only 2*s* and 2*p* electrons, can, at best, form bands of low density and of high energy. Since the metallic bands are only about half-filled, it seems energetically more favourable for the boron atoms to contribute their electrons to these bands.

If three electrons per boron atom participate in the occupation of the *d*-bands, we may expect that the center of gravity of these bands is lowered appreciably because the presence of interstitial borons tends to decrease the potential energy in the region between metal atoms. Because of the directional character of this effect, the *s*-band is probably less affected. This effect, the lowering of the energies of the *d*-bands, both absolutely and with respect to the *s*-band, identifies the band overlap observed in the measurements as occurring between the *d*- and *s*-bands. It also accounts in a natural way for the observed high melting points and cohesive energies. Furthermore, most of the relative physical properties of these compounds should be determined by the degree of filling of the next-higher *s*-band. Since the density of states is low in this band, we expect the cohesive energy to decrease with this filling. This lowering of the melting point in going from column four to column six is actually observed, with the MB_2 phase becoming unstable in column six.

The electrical properties, of course, are in accord with this model. The same model also accounts for the magnetic properties. Our measurements show that all the diborides, with the exception of ZrB_2, show a weak paramagnetism. The susceptibility increases from column four to column five, and is of the right order of magnitude to be due to free electrons, with about one electron per cell in column five.

According to this interpretation, the primary role of the boron atoms is stabilization of a simple hexagonal metallic structure, without any pronounced boron–boron bonds. If boron plays this role, it is easy to understand that the deviations from the ideal boron composition observed in most of these compounds has no critical effect on the lattice stability. Furthermore, one would expect that such interstitial boron tends to make the properties of the material more isotropic than those of the parent metal. This has been confirmed, for instance, for the thermal expansion coefficients in zirconium[14] and ZrB_2.[15]

We can conclude that qualitative understanding of many of the characteristic properties of the transition-metal diborides follows in a natural way from a model based on the assumption that each boron contributes three electrons to the metallic bands. A more quantitative description is not possible until a detailed verification of these suggestions has been carried out for the properties of single crystals.

Acknowledgements—We would like to thank Dr. B. Post for the X-ray analysis of various sumples, and for a number of interesting discussions on the topics of this paper.

REFERENCES

1. Schwarzkopf P. and Kieffer R. *Refractory Hard Metals*. MacMillan, New York (1953).
2. Norton J. T., Blumenthal H. and Sindeband S. J. *Trans. Amer. Inst. Min. (Metall.) Engrs.* **185**, 749 (1949).
3. Volger J. *Phys. Rev.* **79**, 1023 (1950).
4. Landauer R. Private communication.
5. Juretschke H. J., Landauer R. and Swanson J. *J. Appl. Phys.* **27**, 838 (1956).
6. Samsonov G. V. and Neshpor V. S. *Dokl. Akad. Nauk SSSR* **104**, 405 (1956).
7. Samsonov G. V. *Zh. Tekh. Fiz.* **26**, 716 (1956).
8. Wilson A. H. *The Theory of Metals* (2nd Ed.) University Press, Cambridge (1953).
9. Kiessling R. *J. Electrochem. Soc.* **98**, 166 (1951).
10. Post B., Glaser F. W. and Moskowitz D. *Acta Met.* **2**, 20 (1954).
11. Mott N. F. and Jones H. *Properties of Metals and Alloys*. University Press, Oxford (1936).
12. Coles B. R. *Phys. Rev.* **101**, 1254 (1956).
13. Post B. Private communication.
14. Skinner G. B. and Johnston H. L. *J. Chem. Phys.* **21**, 1383 (1953).
15. Bienenstock A. Unpublished M.S. thesis, Polytechnic Institute of Brooklyn (1957).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 128–134.

SINGLE-CRYSTAL ELASTIC CONSTANTS OF INDIUM

D. R. WINDER* **and CHARLES S. SMITH**

Case Institute of Technology, Cleveland, Ohio

(*Received* 27 *June* 1957)

Abstract—The elastic constants of tetragonal indium have been measured by the ultrasonic pulse-echo method. Single crystals were cut so that pulses propagated in the approximate directions [100], [110], and [011] respectively. The nine wave velocities allow the six constants to be determined with three internal checks. If indium is regarded as a tetragonally deformed f.c.c. structure, the constants may be referred to axes coincident with the deformed cubic axes, and in this reference frame the values (expressed in units of 10^{11} dyne cm^{-2}) obtained are:

$$C_{11} = 4{\cdot}45, \qquad C_{33} = 4{\cdot}44,$$
$$(C_{11}-C_{12})/2 = 0{\cdot}253,$$
$$(C_{11}+C_{33})/4-C_{13}/2 = 0{\cdot}197,$$
$$C_{44} = 0{\cdot}655, \qquad C_{66} = 1{\cdot}22.$$

The constants and combinations representing stiffness to shear are low, and are anisotropic in contrast to those for aluminum. The microscopic parameters, particularly those pertaining to overlap electrons, which appear in the calculation of LEIGH for aluminum can be adjusted to account for these features of indium.

INTRODUCTION

THE single-crystal elastic constants of indium have been measured as part of a program aimed at investigating the nature of crystal binding forces. The structure of indium may be most usefully regarded as a tetragonally deformed face-centred-cubic structure with $c/a = 1{\cdot}08$. There are three valence electrons that therefore must overlap into the second Brillouin zone from the first zone, which can contain a maximum of two electrons per atom. As for aluminum[1] and magnesium[2,3] there must be large contributions to the elastic stiffnesses from the Fermi energy of the first zone, of the overlap electrons, and of possible holes in the first zone. VERNON and WEINTROUB[4] in fact have already suggested that a knowledge of the elastic constants would help in understanding the properties of this metal in a paper in which they report that indium has a negative expansion coefficient in the *c*-direction.

The point-group symmetry of indium is $4/mmm$, which means that there are six independent elastic constants. These six have been determined by measuring the velocity of acoustic longitudinal and shear waves in several directions in the crystal by the ultrasonic pulse-echo method.[5] The constants reported in this paper are referred to axes coincident with the deformed cubic axes for ease of interpretation and of comparison with trivalent aluminum.

EXPERIMENTAL PROCEDURE

Ingots were prepared from indium pellets of purity 99·97 per cent as furnished by the Indium Corporation of America. The pellets were cleaned with nitric acid and melted in a graphite crucible into an ingot the shape of a cylinder 1·6 cm in diameter and about 6·4 cm long with a 60° conical end and a tip 0·24 cm in diameter and 0·48 cm long. The ingot was removed, cleaned with nitric acid and packed in a soft mold of Buehler 1557 AB levigated alumina powder contained in a cylindrical graphite cup. The powder was tamped and vibrated into place about the ingot which was centred in the cup. A soft-mold technique was used because the coefficient of thermal expansion in the *c*-direction is negative for

* Now at the National Carbon Research Laboratories, Parma, Ohio.

VOL. 4 1958

indium; if a hard crucible were used and the *c*-direction were other than vertical, the single crystal would expand against the walls and distort upon cooling after solidification. The loaded crucible was mounted in a vertical, resistance-heated tube furnace, and a copper rod, 0·95 cm in diameter and 14·4 cm long, was screwed into the lower end of the cup and into a water-cooled cap at the bottom of the furnace tube. The enforced temperature gradient was enhanced by positioning the crucible in a region of high temperature gradient, so that at the beginning of freezing the temperature gradient across the crucible was about 13°C/cm and at the end of freezing it was about 6°C/cm. The rate of cooling during freezing was about 1°C/min, and the lower end of the charge was superheated to about 85°C above the melting temperature (156°C) in an atmosphere of nitrogen.

The crystal was removed at room temperature, cleaned with nitric acid, etched with Villela's reagent,[6] and examined for grain boundaries and optical flash planes. The optical orientation was accomplished by an optical goniometer consisting of a two-circle crystal mount, a hemicylinder of squared paper, and a spot light whose beam entered through a hole in the paper and was reflected from the crystal back to the paper, thus displaying the symmetry of the crystal. This information was verified by a second examination after re-etching the crystal for 2 min in 12·5 cc HOH, 22·5 cc HCl, 7·5 cc HNO_3, and 6 g HF.

The crystal was scribed to indicate the orientation and was cut accordingly. The crystal was fastened in a Plexiglas miter box with Dow Corning High Vacuum Grease, cooled to the temperature of liquid nitrogen, and maintained there by flooding with liquid nitrogen, while a sample about 1 cm long was cut with a jeweler's saw, using number 0 blades. This cooling was necessary for free, distortionless cutting. To obtain flat parallel ends, the cylindrical sample was lapped in a steel lapping ring on metallographic polishing paper from grade 2 to 2/0, using a lubricant of mineral spirits for the first rough cuts and dry paper for the final paper polishing. To remove cold work, the ends were then electropolished at 1·5 A for 2 min on each end in a solution of 1 part concentrated nitric acid and 2 parts methanol.

An appropriately cut quartz-crystal 10-megacycle transducer was attached to one end of the sample with phenol salicylate. The transit times of acoustic pulses in the crystal were measured by the ultrasonic pulse-echo method in the modification described by Eros[7] for the longitudinal and two shear modes.

Three other measurements are required for the calculation of the elastic constants: the length of the sample, the density ρ, which was determined by hydrostatic weighing to be in agreement with the X-ray valve $\rho = 7{\cdot}30$, and the orientation of the crystal with respect to an end. This orientation of the crystal was accomplished by the back-reflection technique. To obtain a satisfactory Laue pattern, the surface of the sample had to be free of cold-work distortion; also the sample had to be cooled in a liquid-nitrogen bath to reduce the thermal motion of the atoms.

An experiment was performed to assess the magnitude of a systematic error which may occur in the ultrasonic pulse-echo method. This "transit-time error" arises in the acoustic reflections at the quartz-cement-specimen interfaces.[8] The previously described[9] "dummy quartz" technique was carried through, with the result that the corrections turned out to be smaller than the other experimental uncertainties. Consequently no transit-time correction was made.

A sample from one of the crystals was analyzed

Table 1. *The direction cosines of the direction of propagation, the approximate directions of particle motion, and the values of* ρV^2 *in units of* 10^{11} *dyne* cm^{-2} *for each crystal*

Crystal	Direction cosines ($l\ m\ n$)	Approximate direction of particle motion	ρV^2 values
a	(0·986, 0·155, 0·0539)	100	4·56
		010	1·15
		001	0·652
b	(0·710, 0·705, 0·0323)	110	5·43
		1$\bar{1}$0	0·254
		001	0·653
c	(0·0348, 0·717, 0·698)	011	4·93
		100	0·926
		0$\bar{1}$1	0·202
d	(0·111, 0·699, 0·735)	011	4·92
		100	0·869
		0$\bar{1}$1	0·217

spectroscopically by the Herron Testing Laboratories, Inc., of Cleveland, Ohio, for the following impurities: Pb, Sn, Fe, Zn, Cd, Sb, As, Ag, Cu, Al, and Bi. The impurities present in the sample were 0·01 per cent iron and 0·01 per cent copper.

RESULTS

The primary data, consisting of values of ρV^2 and of direction cosines of the propagation directions are listed in Table 1. These data were reduced to elastic stiffnesses by the technique described fully by NEIGHBOURS.[10] The procedure used may be described with the aid of Table 1 and Fig. 1, in

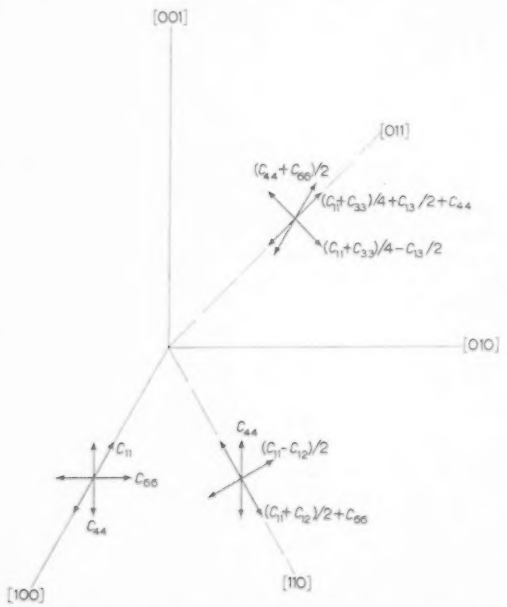


FIG. 1. Combinations of elastic constants given by ρV^2 for the propagation directions and polarizations shown.

which the combinations of constants corresponding to ρV^2 for the ideal directions of propagation, [100], [110], and [011], are shown. Zeroth order values for C_{11}, C_{44}, and C_{66} are obtained directly from the ρV^2 values for crystal a, for $(C_{11}-C_{12})/2$ from crystal b, and for $(C_{11}+C_{33})/4-C_{13}/2$ and $(C_{11}+C_{33})+C_{13}/2+C_{44}$ from crystals c and d. First-order constants were then computed by applying small-angle corrections. The computation then proceeded by iteration to a self-consistent result. Perturbation terms[10] were included, although these are small because the shear constants of indium are small, as is discussed further on.

From the four crystals there are twelve measurements of ρV^2 connecting the six elastic constants. Thus six internal checks are possible and these are shown in Table 2. These checks are divided into

Table 2. The six internal checks on the elastic constants, in units of 10^{11} *dyne cm*$^{-2}$*, as determined from measurements on the indicated crystals* a, b, c, *and* d

Constant	Values	
	a	b
C_{44}	0·659	0·651
C_{66}	1·27	1·21
	a	c
$(C_{44}+C_{66})/2$	0·963	0·921
	c	d
C_{33}	4·44	4·42
$(C_{44}+C_{66})/2$	0·921	0·925
$(C_{11}+C_{33})/4-C_{13}/2$	0·197	0·205

three checks on the important shears and three between crystals c and d which have roughly the same orientation.

Finally the adiabatic and isothermal constants and compliances for tetragonal indium are shown

Table 3. The values at 298°K *for indium of the measured adiabatic elastic constants,* C^S*, the computed adiabatic compliances,* S^S*, the isothermal compliances,* S^T*, and the isothermal constants* C^T*. Units are* 10^{11} *dyne cm*$^{-2}$ *for C and* 10^{-11} *cm*^{2} *dyne*$^{-1}$ *for S.*

Subscript	C^S	S^S	S^T	C^T
11	4·45	1·494	1·498	4·21
12	3·95	−0·506	−0·502	3·71
13	4·05	−0·902	−0·902	3·82
33	4·44	1·870	1·870	4·22
44	0·655	1·527	1·527	0·655
66	1·22	0·820	0·820	1·22

in Table 3. The conversions have been made as previously described for hexagonal magnesium[2] with the addition of the relation $S_{66} = C^{-1}{}_{66}$. For the conversion from adiabatic to isothermal conditions, the following data have been used: Temperature 298°K, specific heat $Cp = 6{\cdot}56$ cal mole^{-1} deg^{-1}, thermal expansion coefficients $\gamma_1 = \gamma_2 = 49 \times 10^{-6}$ deg^{-1}, $\gamma_3 = -6{\cdot}5 \times 10^{-6}$ deg^{-1} The latter are the results of VERNON and WEINTROUB,[4] extrapolated slightly.

DISCUSSION OF RESULTS

The internal checks shown in Table 2 suggest an uncertainty of 5 per cent in the elastic constants. The results quoted in Table 3 are somewhat better than this in some respects however. C_{44} is directly measured and is internally checked; an uncertainty of 1 per cent is therefore assigned. C_{11} is directly measured and is felt to be known to 2 per cent. All four crystals yield values for $C_{44}+C_{66}$, and Table 2 shows that the result for C_{66} from crystal a is the source of the largest discrepancy. Since this measurement was also the poorest of the set, crystal a has been weighted lightly in the final result for C_{66}, which is assigned an uncertainty of 2 per cent. The remaining elastic constants, C_{33}, C_{12} and C_{13}, are all arrived at indirectly, and an uncertainty of 5 per cent is therefore assigned; certain combinations mentioned below are, however, known better than this.

The interpretation of elastic-constant data for a polyvalent metal such as indium is usually confined to constants and combinations of constants which represent shear stiffnesses. In Table 4 are shown such stiffnesses, and it may be remarked that these values are quite well known since the experiment has been designed to measure them as directly as possible. C_{44} and C_{66} are pure shear constants measuring the stress/strain ratio for a change in angle between the a–c and a–a axes respectively. $C' = (C_{11}-C_{12})/2$ is a pure shear constant measuring the stiffness to a change in angle between ⟨110⟩ axes. $C'' = (C_{11}+C_{33}-2\,C_{13})/4$ gives the shear-stress/shear-strain ratio for a shear strain which changes the angle between orthogonal ⟨011⟩ axes; this constant is not a pure shear constant however. Closely related to C'' is the combination $(C_{11}+2\,C_{33}+C_{12}-4\,C_{13})$, which gives the second derivative of the crystal energy with respect to ϵ for a strain system such that unit distance along the c-axis becomes $(1+\epsilon)$ and unit distances along the a-axis become $(1+\epsilon)^{-\frac{1}{2}}$. This combination is basic to some future explanation of the tetragonality of indium. Finally, in Table 4, are given combinations which we shall call "pseudo-cubic shear constants", that is, the weighted averages $(2C_{44}+C_{66})/3$, and $(2C''+C')/3$. The latter will be used in the next section.

Table 4. Combinations of constants representing the stiffness of indium to shear. Units are 10^{11} *dyne cm*$^{-2}$ *and the values are isothermal*

Constant	Value
C_{66}	1·22
$C' = (C_{11}-C_{12})/2$	0·253
C_{44}	0·655
$C'' = (C_{11}+C_{33}-2\,C_{13})/4$	0·197
$(C_{11}+C_{12}+2\,C_{33}-4\,C_{13})$	1·08
$(2\,C_{44}+C_{66})/3$	0·843
$(2\,C''+C')/3$	0·215

In connection with the shear stiffness results, several points are of interest. First, the values are low for a metal relative to the "normal stress" constants C_{11} and C_{33}, and are also low compared with the corresponding shear constants of aluminum, which has the same structure and valence. Second, the anisotropy ratios C_{66}/C' and C_{44}/C'' are quite normal for a metal, in contrast with aluminum, which is anomalously isotropic. Taken together these points can be put another way, i.e. the constants C' and C'' are extraordinarily low. This last fact is seen immediately to be consistent with the barely tetragonal structure of indium, since these constants measure the stiffness to strains which change the crystal axial ratios.

Finally we comment that, as a consequence of the first point made above, the normal stress constant C'_{11} is nearly isotropic in spite of the substantial shear anisotropy. Equations for C'_{11} in the a–a plane ($n = 0$) and the a–c plane ($l = 0$) are as follows

$$n = 0, \qquad C'_{11} = C_{11} - 4l^2m^2(C' - C_{66})$$

$$l = 0, \qquad C'_{11} = C_{11}m^2 + C_{33}m^2 - 4m^2n^2(C'' - C_{44}).$$

Fig. 2 shows the isotropy of C'_{11}, the relative magnitudes of C'_{11} and the shear constants, and the anisotropy of the latter.

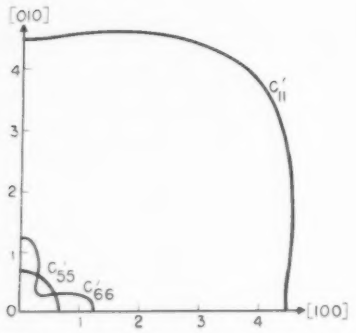


FIG. 2. Polar plots showing the nearly isotropic longitudinal constant C'_{11} and the anisotropic shear constants C'_{55} and C'_{66}. Units of 10^{11} dyne cm^{-2}.

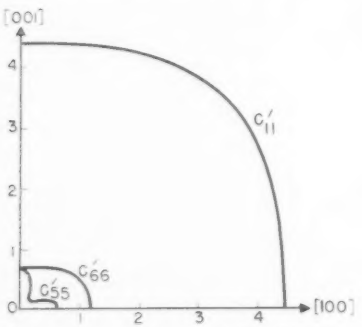


FIG. 2*a* shows the *a*–*a* plane and Fig. 2*b* the *a*–*c* plane.

INTERPRETATION

It is possible to make a preliminary and qualitative interpretation in microscopic terms of the shear-constant data for indium by modifying suitably the theory of LEIGH[1] for f.c.c. aluminum. For this purpose we consider indium to be a pseupo-cubic fcc structure with lattice parameter $a' = (a^2c)^{\frac{1}{3}} = 4{\cdot}710$ Å as compared with $a = 4{\cdot}049$ Å for aluminum. The theory will then apply to the indium pseudo-cubic shear constants $(2C_{44}+C_{66})/3$ and $(2C''+C')/3$. In the same spirit we shall deal with the usual first Brillouin zone for a fcc structure, bounded by six square faces normal to $\langle 100 \rangle$ and eight hexagonal faces normal to $\langle 111 \rangle$, just as for aluminum except for size. Since there are three electrons per atom, we shall suppose the first Brillouin zone to be completely filled with two electrons, the remaining electron overlapping into the second zone, as LEIGH supposed for aluminum.

LEIGH calculated for such a structure the following contributions to the shear stiffnesses: (1) The "Coulomb" term, arising from the electrostatic interaction between positive ion cores and valence electrons. This term contains an adjustable effective charge parameter Z^2. (2) The "first zone" term, arising in the Fermi energy of the electrons in the filled first Brillouin zone. There is here an adjustable parameter α_0 which is the reciprocal effective mass ratio for these electrons (3) Terms arising from the Fermi energy of hexagonal face overlap electrons and from the transfer of these electrons between Brillouin zone faces during shear. There are n_h of these electrons per atom per pair of hexagonal faces, and the number of states per atom per unit energy range per pair of hexagonal faces is designated by N_h. The electron energy at the center of the hexagonal face, and in the second Brillouin zone, is indicated by E_h. (4) Similar terms for square face overlap electrons with parameters n_s, N_s, and E_s.

Table 5 shows the theoretical expressions for these contributions to the pseudo-cubic shear constants. The table is given in terms of ΩC, where Ω is the atomic volume. The observed values from this paper and the numerical constants have been expressed in electron-volts, with the latter taken from LEIGH's paper and scaled to the

Table 5. The theoretical terms contributing to the pseudo-cubic shear constants of indium. Units are electron-volts

Term	$\Omega(2C''+C')/3$	$\Omega(2C_{44}+C_{66})/3$
Coulomb	$0{\cdot}163\ Z^2$	$1{\cdot}461Z^2$
First zone electrons	$1{\cdot}15\alpha_0$	$3{\cdot}44\alpha_0$
Hexagonal-face overlap electrons	$\frac{8}{3}n_hE_h$	$\frac{8}{3}n_hE_h - \frac{1.6}{0}N_hE_h^2$
Square-face overlap electrons	$2n_sE_s - 2N_sE_s^2$	$2n_sE_s$
Observed	0·35	1·38

pseudo-cubic lattice parameter a'. It is now required that the theoretical terms add up to give the observed pseudo-cubic shear constants, which requirement yields two equations. In addition there are two equations expressing normalization conditions. For one overlap electron per atom we have

$$4n_h+3n_s = 1.$$

Also the total density of states, N, is known from electronic specific heat data, so that

$$4N_h+3N_h = N = 0{\cdot}768,$$

where we have used the specific-heat value quoted by KEESOM and PEARLMAN.[11] There are no other directly useful data for indium to our knowledge; in particular the Fermi level is not known from soft X-ray spectroscopy, as it is for aluminum.

There are then four equations in eight unknowns. However, the situation is not so completely open as this statement would seem to imply, because all unknowns must be positive or zero and other conditions of reasonableness apply to them. We start by fixing $E_h = 6{\cdot}70$ eV, and $E_s = 8{\cdot}28$ eV.

These numbers are scaled from the corresponding values assigned by LEIGH for aluminum and are somewhat greater than the respective values of $\hbar^2k^2/2m$, as they should be. Similarly α_0 is placed at 1·00, as for aluminum. At this point, inspection of Table 5 and recognition that one expects $N_s = \text{const.}\ n_s^{\frac{1}{2}}$ allow the conclusion that square-face overlap electrons must be present in order to account for the observed value of $\Omega(2C''+C')/3$. This conclusion was also reached by LEIGH for aluminum. The conclusion holds for indium even for any small, but not unreasonably small, values of Z^2, α_0, and E_h which might be assigned, since any one of the first three terms would be as large individually as the observed value of this elastic constant.

The next step is to fix a value for the Coulomb parameter Z^2. Recognizing that because of the necessary relation $E_h < E_s$ it will follow that $n_h > n_s$, that $n_s > 0$, and that we shall expect $N_s \sim n_s^{\frac{1}{2}}$, the four equations in five unknowns (E_h, E_s, α_0 fixed) may be inspected for the range of Z^2 which yields possible solutions. This range turns out to be quite narrow, i.e. $4{\cdot}25 > Z^2 > 3{\cdot}25$. The range of N_s which will yield possible solutions is also quite narrow, $0{\cdot}05 > N_s > 0{\cdot}04$, so that the limit $Z^2 = 4{\cdot}25$ actually corresponds to $n_s = 0$, while the limit $Z^2 = 3{\cdot}25$ corresponds to the maximum value of $n_s = 0{\cdot}333$. The value $Z^2 = 4{\cdot}10$ has therefore been chosen. Having set four parameters, the remaining four may be deduced from the equations. The parameters found in this way are given in Table 6, along with the numerical values of the various contributions to the elastic constants. It is seen that it is possible to account for the observations with the theory; even this much is by no means certain in advance when the various terms have such very different anisotropies.

Furthermore the parameters which have been used are all quite normal ones with the exception of Z^2, which is discussed below. The density of states effective mass of the square-face overlap electrons $(m^*/m)_s$, computed from $N = \text{const.}\ (m^*/m)n^{\frac{1}{2}}$, comes to $(m^*/m)_s = 0{\cdot}35$. Undoubtedly

Table 6. Numerical values of the theoretical terms contributing to the pseudo-cubic shear constants of indium together with the corresponding values of microscopic parameters. Energy units are eV; other units are given in the text

Terms	Parameters	$\Omega(2C''+C')/3$	$\Omega(2C_{44}+C_{66})/3$
Coulomb	$Z^2 = 4{\cdot}10$	0·67	5·99
First zone electrons	$\alpha_0 = 1{\cdot}00$	1·15	3·44
Hexagonal-face overlaps	$E_h = 6{\cdot}70$, $n_h = 0{\cdot}219$, $N_h = 0{\cdot}159$	3·91	−8·74
Square-face overlaps	$E_s = 8{\cdot}28$, $n_s = 0{\cdot}0418$, $N_s = 0{\cdot}0442$	−5·38	0·69
Total		0·35	1·38
Observed		0·35	1·38

these electrons have anisotropic effective mass, and if the components in the Brillouin zone plane are placed at unity, the component normal to the Brillouin zone plane is $(0{\cdot}35)^3$. These values are reasonable and so is the value of $(m^*/m)_h = 0{\cdot}73$ computed from the above formula for the hexagonal-face overlaps. The effective masses are also similar to those for aluminum.

The value $Z^2 = 4{\cdot}10$ of the Coulomb parameter seems at first somewhat low. The physical limits of Z^2 for trivalent metals are $9 > Z^2 > 0$, and LEIGH felt that, for aluminum, $Z^2 = 7$ was the lowest reasonable value. As has been seen however, $Z^2 = 4{\cdot}10$ is more or less forced by the experimental data. The value of Z^2 can be increased by reducing the chosen value of α_0, but only slightly so, as a detailed examination of the equations shows. The reason for this situation is that the anisotropies of the Coulomb and first-zone terms are so different. By way of justification it may be pointed out that Z^2 for indium may well be expected to be smaller than Z^2 for aluminum, because of the higher atomic number of the former. By way of precedent, practically the same value of $(Z^2/Z^2\text{max})$ has been proposed in a similar treatment of magnesium.[3]

A colleague of the writers has under way a generalization of the theory for the detailed tetragonal shear constants C_{44}, C_{66}, C', and C'', so the present writers will go no further than to point out a few qualitative features. Each physical term in the cubic theory will be modified in the tetragonal case; degeneracies will be removed so to speak. In the tetragonal form $n_s(c)$ will be expected to be greater than $n_s(a)$, since the Brillouin zone plane normal to the c-axis will have moved inward in reciprocal space. Square-faced overlap transfer terms should then be greater for c–a transfer than for a–a transfer and so on, on this account alone, we should expect $C' > C''$, as observed. A shear strain that "makes" indium tetragonal is of the C' or C'' type, and it may well be that it is the reduction of energy arising in these transfer terms which is responsible for the tetragonal structure itself. Here, however, competing effects have been ignored, e.g. those which keep aluminum cubic.

Acknowledgements—This research was supported by the Office of Naval Research and the National Carbon Company. Our Colleague, J. R. REITZ, has helped greatly with the interpretation of the results. We are indebted to A. HRUSCHKA for the support of his excellent machine-shop.

VOL. 4 195

REFERENCES

1. LEIGH R. S. *Phil. Mag.* **42,** 139 (1951).
2. LONG T. R. and SMITH C. S. *Acta Met.* **5,** 200 (1957).
3. REITZ J. R. and SMITH, CHARLES S. *Phys. Rev.* **104,** 1253 (1956).
4. VERNON E. V. and WEINTROUB S. *Proc. Phys. Soc.* B **66,** 887 (1953).
5. HUNTINGTON H. B. *Phys. Rev.* **72,** 321 (1947).
6. CARAPELLA S. C. and PERETTI E. A. *Metal Progr.* **56,** 666 (1949).
7. EROS S. (private communication.)
8. NEIGHBOURS J. R., BRATTEN F. W. and SMITH, CHARLES S. *J. Appl. Phys.* **23,** 389 (1952).
9. SMITH, CHARLES S. and BURNS J. W. *J. Appl. Phys.* **24,** 15 (1953).
10. NEIGHBOURS J. R. *J. Acoust. Soc. Amer.* **26,** 865 (1954).
11. KEESOM P. H. and PEARLMAN N. *Encyclopedia of Physics* Vol. 14, p. 282. Springer, Berlin (1956).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 135–143.

THERMODYNAMICS AND MAGNETIC STRUCTURES OF THE ALLOTROPIC MODIFICATIONS OF MANGANESE

R. J. WEISS* and K. J. TAUER†‡

*Crystallographic Laboratory, Cavendish Laboratory, Cambridge, England

†Chemistry Department, Boston College, Massachusetts

(*Received* 18 *June* 1957)

Abstract—The free energy, entropy and enthalpy of the allotropic modifications of Mn are separated into their magnetic, electronic and lattice components. The energy difference at absolute zero of the various modifications are determined and it is shown that similar to Fe the high temperature body centred modification would be stable again at low temperature if not for the presence of the complicated α- and β-phases. The thermodynamic parameters are consistent with recent ideas on the electronic structure of transition metals and some suggestions are made concerning the presence of these complicated phases. In particular the α-, β- and δ-phases appear related electronically.

INTRODUCTION

A PREVIOUS analysis of the allotropic modifications of iron[1,2] have shown that the free energy is separable into magnetic, lattice and electronic components, the magnetic component representing the disordering of the magnetic spin system, the lattice component the excitation of atomic oscillators and the electronic component the excitation of electrons at the Fermi surface. It is the purpose of this note to perform a similar analysis on manganese and to show that the

‡ Both authors are associated with the Materials Research Laboratory, Watertown, Massachusetts, U.S.A.

Table 1. *Atoms per unit cell; lattice parameter, a; atomic volume; stability range; magnetic structure; magnetic critical temperature; and Bohr magneton number for each equivalent lattice site; all for the four modifications of Mn*

Modification	Atoms Unit cell	(a) 300°K	Atomic volume 300°K(A^3)	Stability range	Magnetic structure
(cubic) α	58	8·894	12·09	0–1000°K	antiferro
(cubic) β	20	6·300	12·50	1000–1364°K	para ?
(f.c.c.)*	4	3·715	12·80	1364–1410°K	antiferro
(b.c.c.)	2	2·978	13·15	1410–1450°K	antiferro

Modification	Magnetic critical temp. (°K)	Bohr magneton number
(cubic) α	95°K	I–1·5 II–1·5 III–1·5 IV–0
(cubic) β	~0 ?	I–1·0 II–1·0 ?
(f.c.c.)*	580°K	2·3
(b.c.c.)	627°K ?	~1

* Tetragonal at room temperature $c/a = 0{\cdot}950$.

thermodynamic properties of manganese are consistent with some recent ideas[3] on the electronic structures of transition metals.

Manganese crystallizes in four allotropic modifications whose properties are summarized in Table 1.

The experimental data used for this analysis are tabulated in Table 2.

Table 2. Temperatures at which experimental specific heat, neutron diffraction and heat content data were used for thermodynamic analyses of Mn. Numbers in parenthesis are references

Modification	Specific Heat (°K)	Neutron Diffraction (°K)	Heat Content (°K)
α	12–20[4] 50–300[5] 300–1000[6]	4,300[7]	300–1000[8]
β	4–20[10] 12–20[4] 1000–1050[6]	4,300[7]	1000–1364[8]
γ	50–300[5]	300–700[9]	1364–1410[8]
δ	—	—	1410–1450[8]

αMn

The separation has been performed for αMn in the range 12–20°K and 50–300°K and is described in this journal.[11] The separated magnetic entropy and energy agree with neutron diffraction determinations of the magnetic structure. The parameters deduced for this separation have been applied to evaluating the heat content and specific heat of αMn from 300°K to 1000°K and compare very well with the data of NAYLOR[8] and ARMSTRONG[6] after making a slight change of the elec-

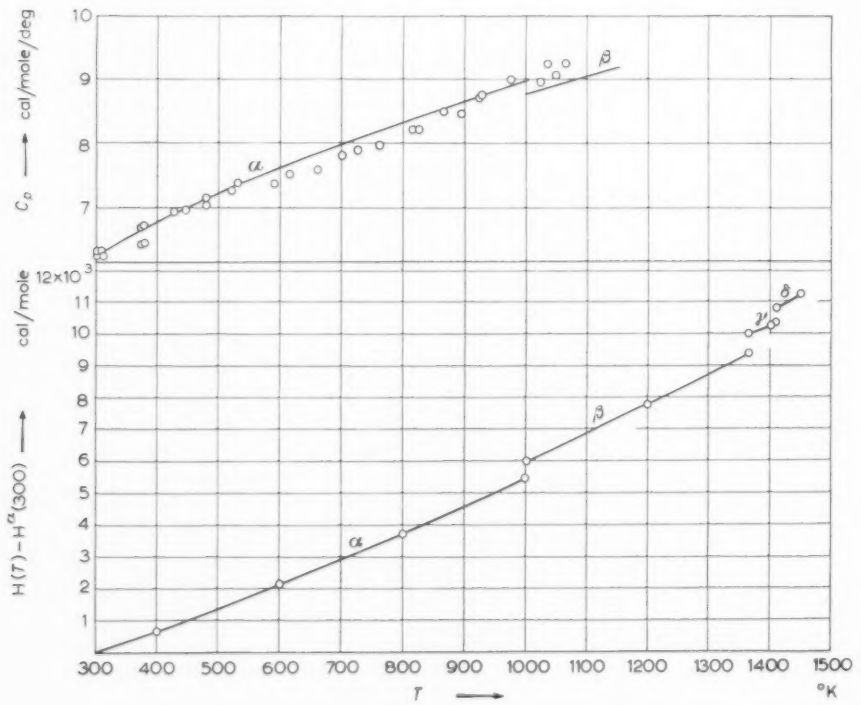


FIG. 1. Calculated (solid lines) and observed (circles) specific heat (upper) and heat content (lower) curves for Mn above room temperature.

Table 3. Debye temperature, θ; electronic specific heat coefficient, γ; total magnetic entropy, S_{Mag}; magnetic critical temperature, T_c°K; total magnetic enthalpy, H_{Mag}; total energy difference at absolute zero relative to αMn; lattice energy difference at absolute zero relative to αMn; temperature of the phase transition; and latent heat of the transition; all for the four modifications of Mn

Modification	θ°K	$\gamma \times 10^4$	S_{Mag}	T_c°K	H_{Mag}	$(H_0-H_0^{\alpha})_{Total}$ cal/mole
α	460	25·3	1·071	95	105	—
β	422	22·5	1·38	~0	~0	724·5
γ	355	11·2	2·35	580*	1210	1137·8
δ	370	22·5	1·38	627	990	863·5

Modification	$(H_0-H_0^{\alpha})^{\dagger}_{Lattice}$	Transition temp. (°K)	Latent heat cal/mole
α	≡ 0	—	—
β	704·5	1000	543 ±50
γ	2476·8	1364	572 ±50
δ	1948·5	1410	472 ±50

* 640°K for f.c.t.

† Zero point correction made = $(9/8)R\theta$.

Table 4. The magnetic, electronic and lattice components of entropy and enthalpy for αMn as a function of temperature

T(°K)	S^{α}_{Mag}	$\left\{ H^{\alpha}_{Mag} - H_0{}^{\alpha}_{Mag} \right\}$	S^{α}_{Elect}	$\left\{ H^{\alpha}_{Elect} - H_0{}^{\alpha}_{Elect} \right\}$	$S^{\alpha}_{Lattice}$	$\left\{ H^{\alpha}_{Lattice} - H_0{}^{\alpha}_{Lattice} \right\}$
298·16	1·0714	105	0·755	112	5·8250	978
400	1·0714	105	1·012	202	7·4493	1537
600	1·0714	105	1·518	456	9·8110	2726
800	1·0714	105	2·024	810	11·6659	3984
1000	1·0714	105	2·530	1265	13·0990	5279
1200	1·0714	105	3·036	1820	14·2891	6579
1400	1·0714	105	3·542	2480	15·3213	7947
1600	1·0714	105	4·048	3239	16·2254	9247

tronic specific heat coefficient, γ, from 28×10^{-4} cal/mole/deg^2 to $25{\cdot}3 \times 10^{-4}$ cal/mole/deg^2 a change within the error of the low temperature measurements. The resulting parameters are given in Table 3 and the calculated and observed specific heat and heat content curves given in Fig. 1. In addition Table 4 gives the resulting magnetic, electronic and lattice entropies and enthalpies from which the free energy can be evaluated. All the calculated thermodynamic data for αMn are in agreement with the experimentally observed specific heat, heat content and neutron diffraction data over all temperature ranges.

βMn

KASPER and ROBERTS have reported that βMn is not ferro- or antiferro-magnetic down to 4°K (magnetic measurements and neutron diffraction). The specific heat data of HOARE 12–20°K and SMITH* 1–20°K are both anomalous, the former in its failure to separate into T and T^3 terms and the latter in evidencing an unusually high γ (~ 124 cal/mole/deg). Both these experimenters report identical results for αMn and so it is concluded that the anomalous specific heat is due to small

* Private communication.

differences in sample impurity (αMn for example becomes ferrimagnetic at low temperatures due to occluded H). The high temperature heat content data of NAYLOR and the specific heat data of ARMSTRONG are unable to give information concerning the Debye θ or the magnetic structure since the temperatures are well above the Debye temperature and magnetic critical temperature. However both these results are able to give values of the electronic specific heat coefficient, γ, the former yielding $\gamma = 23{\cdot}8\times10^{-4}$ cal/mole/deg^2 and the latter $\gamma = 21{\cdot}1\times10^{-4}$ cal/mole/deg^2. The average of these two ($22{\cdot}5\times10^{-4}$) has been chosen for our analysis. With this average value of γ we can choose combinations of Debye θ and magnetic entropy to yield the total entropy of βMn at the transition temperature 1000°K, since $S^{\beta}(1000°\text{K}) = S^{\alpha}(1000°\text{K}) + \Delta H/1000 = 17{\cdot}2419$ e.u. where ΔH is the latent heat measured by NAYLOR. These combinations are given in Table 5.

Table 5. Combination of Debye θ and magnetic entropy which yield the correct total entropy of β-manganese at its transition temperature (1000°K) after subtracting the electronic entropy.

θ(°K)	S_{Mag} e.u.
336	0
360	0·387
380	0·631
400	1·054
420	1·349
400	1·623
460	1·886

While any specific choice of the above will make little difference in the high temperature thermodynamic functions, we can choose a set from the following considerations.

(1) Independent of the choice of θ there appears to be a large additional entropy in the data of HOARE and SMITH and in both cases this additional entropy is still increasing at 20°K. This additional entropy ($> 0{\cdot}5$ cal/mole/deg.) could arise from some short range spin coupling very sensitive to impurity content. KASPER has pointed out that one of the two crystallographically equivalent sites are geometrically unable to give rise to a consistent long range antiferromagnetic coupling, hence large moments $\sim 1\mu_B$ may not lead to superlattice lines.*

(2) The Debye θ of the β-phase is probably not too different from that of the α-phase.

We have therefore chosen a $\bar{\mu}_B \cong 1$ giving rise to $S^{\beta}\text{mag} = 1{\cdot}38$ and $\theta = 422°\text{K}$. The thermodynamic functions are thus given in Table 6 and the good agreement with the heat content data of NAYLOR and specific heat data of ARMSTRONG is shown in Fig. 1.

γMn

Neutron diffraction data on Mn-rich copper manganese alloys[9] extrapolated to pure Mn yield

* Susceptibility of βMn is currently being measured in this laboratory from 77°K to room temperature.

Table 6. The magnetic, electronic and lattice components of entropy and enthalpy for βMn as a function of temperature

T	S^{β}_{Mag}	$(H^{\beta}_{\text{Mag}}-H_0^{\beta})$	S^{β}_{Elect}	$(H^{\beta}_{\text{Elect}}-H_0^{\beta})$	$S^{\beta}_{\text{Lattice}}$	$(H^{\beta}-H_0^{\beta})_{\text{Lattice}}$
298·16	1·38	0	0·671	99·5	6·2689	1021
400	1·38	0	0·900	180	7·9502	1601
600	1·38	0	1·350	450	10·3827	2808
800	1·38	0	1·800	720	12·1851	4061
1000	1·38	0	2·250	1125	13·6119	5343
1200	1·38	0	2·700	1620	14·8051	6659
1364	1·38	0	3·070	2093	15·6695	7763
1390	1·38	0	3·130	2170	15·8017	7934
1400	1·38	0	3·150	2205	15·8322	8009
1500	1·38	0	3·375	2530	16·3157	8696
1600	1·38	0	3·600	2875	16·7480	9357

the magnetic properties in Table 1. The specific heat data on pure γMn from 50–300°K[5] can be separated into its magnetic, lattice and electronic components. (Neutron diffraction data has not as yet been performed on pure γMn.) In this range the specific heat is given by

$$C_p = C_v(\theta/T)\{1+AT\}+\gamma T+BT^3 \qquad (1)$$

where γ is the electronic specific heat coefficient, B the spin wave coefficient, $C_v(\theta/T)$ the Debye specific heat and $(1+AT)$ the coefficient which corrects C_v to C_p. A is taken to be 10^{-4}/deg. as this value was found to yield good results for αMn (the coefficient of expansion of α, β, γ and δMn are all approximately equal). By evaluating $C_v(\theta/T)\{1+AT\}$ for various values of θ, subtracting this from C_p (observed) and plotting the residual specific heat divided by T versus T^2 we have

$$\frac{C_p(\text{obs.})-C_v(\theta/T)\{1+AT\}}{T}=\gamma+BT^2. \qquad (2)$$

That value of θ which most nearly yielded a straight line was selected and the slope and intercept at $T=0$ gave values for γ and B. This is shown in Fig. 2 and we obtained $\theta = 335°\text{K}$, $\gamma = 11{\cdot}2\times10^{-4}$ cal/mole/deg^2 and $B = 5{\cdot}14/T_c)^3$ cal/mole/deg^4. These values are further substantiated by giving good agreement with the entropy at the β–γ transition (1364°K) where $S^\gamma = S^\beta+\Delta H/1364 =$ 20·5391 e.u. (ΔH is the latent heat measured by NAYLOR).

The low temperature form of γMn is face centred tetragonal due to its magnetic structure. The nearest neighbor interaction is antiferromagnetic but since nearest neighbors are nearest neighbors of each other one cannot arrange a long range order such that each atom is surrounded entirely by anti-parallel spins. The lowest energy long range order state is one in which eight of the twelve nearest neighbors are anti-parallel and four are parallel. The tetragonal distortion is such that the eight anti-parallel nearest neighbor distances are shortened in comparison with the four parallel nearest neighbor distances. This lowers the magnetic energy (at the expense of some lattice

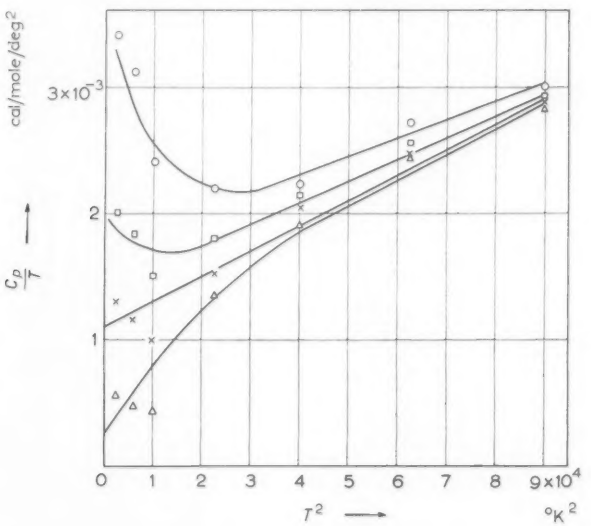


FIG. 2. Analysis of the specific heat of γMn into a Debye and spin wave (T^3) term after subtracting the electronic term. The value of $\theta = 355$ was chosen as most nearly satisfying this separation.

$\bigcirc\theta = 370$ $\times\theta = 355$

$\square\theta = 360$ $\triangle\theta = 350$

Table 7. See Table 6 except substitute γ for β.

T	S^{γ}_{Mag}	$(H^{\gamma}-H_0^{\gamma})_{\text{Mag}}$	$S^{\gamma}_{\text{Elect}}$	$(H^{\gamma}-H_0^{\gamma})_{\text{Elect}}$	$S^{\gamma}_{\text{Lattice}}$	$(H^{\gamma}-H_0^{\gamma})_{\text{Lattice}}$
298·16	0·284	63	0·334	49	7·2181	1125
400	0·689	206	0·448	90	8·9345	1725
600	2·097	918	0·672	201	11·3981	2937
800	2·213	991	0·896	358	13·2096	4196
1000	2·262	1032	1·120	560	14·6171	5486
1200	2·289	1063	1·344	806	15·8429	6808
1364	2·303	1080	1·529	1042	16·7071	7843
1400	2·305	1084	1·568	1098	16·8674	8128
1410	2·306	1086	1·579	1112	16·9215	8212
1500	2·311	1092	1·680	1260	17·3527	8830
1600	2·316	1100	1·792	1434	17·7857	9504

distortion energy) and raises the Néel temperature, compared to that for a face-centred cubic lattice. The high temperature thermodynamic properties are calculated for the face-centred cubic lattice differing only from the tetragonal lattice by a small change in the magnetic properties. The tetragonal to cubic transition appears as a first order phase change in the vicinity of the Néel temperature (no free energy change) and is explicable in terms of order disorder theory as developed for its alloy counter part CuAu.[12]

The separated entropies and enthalpies are given in Table 7. The magnetic free energy was evaluated using the following empirical expressions derived in the appendix. These are probably good to 5 per cent and offer a good approximation in the absence of explicit specific heat data.

$$G_{\text{Mag}} = -RT_c(T/T_c{}^4[0{\cdot}4\ln(2\bar{s}+1)-0{\cdot}2\bar{s}^{\frac{1}{2}}] \quad (3)$$

$$G_{\text{Mag}} = -RT_c\bar{s}^{\frac{1}{2}}\left\{T/T_c\frac{\ln(2\bar{s}+1)}{\bar{s}^{\frac{1}{2}}}-1+\right.$$
$$\left.+T_c/T\left[0{\cdot}8-\frac{0{\cdot}6\ln(2\bar{s}+1)}{\bar{s}^{\frac{1}{2}}}\right]\right\}\ T>T_c \quad (4)$$

where T_c is thecritical temperature, $\bar{s}$ the average spin magnitude and R the gas constant.

The calculated heat content for γMn is given in Fig. 1 and agrees with that of NAYLOR.

δMn

The only data available on δMn is the heat content data of NAYLOR from 1410 to 1450°K plus the latent heat of the γ–δ transition. From the temperature dependence of the heat content data γ is estimated to be $\sim 25\times10^{-4}$ cal/mole/deg² with an error of about 5 cal/mole/deg²$\times10^{-4}$. From the latent heat and the entropy of the γ phase at the transition temperature we have

$$S^{\delta}(1410) = S^{\gamma}(1410)+\frac{\Delta H}{1410} = 21{\cdot}1418 \text{ e.u.} \quad (5)$$

From recent neutron diffraction experiments of KASPER on CrMn (b.c.c.) in which this alloy was found to be antiferromagnetic with an average Bohr magneton number of $\sim 0{\cdot}7$ and from experiments on pure b.c.c. Cr $\mu_B \cong 0{\cdot}4$ we estimate that pure b.c.c. Mn is antiferromagnetic with $\mu_B \cong 1{\cdot}0$. This value is consistent with the interpolated number of atomic d electrons between Cr and Fe (b.c.c.) determined from recent X-ray measurements of electron charge density. This gives

Table 8. Combinations of Debye θ and electronic specific heat coefficient, γ, yielding the correct total entropy of δMn at the transition temperature 1410°K after subtracting the estimated magnetic entropy

θ °K	$\gamma(\times10^4)$ cal/mole/deg²
460	31·5
440	29·7
420	27·9
400	25·9
380	23·6
370	22·5
350	19·9

Table 9. Same as Table 6 except for δMn

T(°K)	S^{δ}_{Mag}	$(H^{\delta}-H_0^{\delta})_{\text{Mag}}$	$S^{\delta}_{\text{Elect}}$	$(H^{\delta}-H_0^{\delta})_{\text{Elect}}$	$S^{\delta}_{\text{Lattice}}$	$(H^{\delta}-H_0^{\delta})_{\text{Lattice}}$
298·16	0·102	23	0·672	100	6·9866	1097
400	0·246	73	0·900	180	8·6934	1693
600	0·830	371	1·350	405	11·1531	2917
800	1·113	564	1·800	720	12·9572	4164
1000	1·209	648	2·250	1125	14·3992	5457
1200	1·262	707	2·700	1620	15·6025	6774
1400	1·293	747	3·150	2204	16·6368	8114
1410	1·294	748	3·1725	2238	16·6753	8171
1500	1·304	763	3·375	2535	17·1043	8806
1600	1·313	786	3·600	2780	17·5464	9473

$T_c \cong 627$°K with eight nearest neighbors antiparallel. At $T = 1410$°K we have $S_{\text{Mag}} = 1{\cdot}294$ e.u. and the following combinations of θ and γ yield the total entropy at the transition temperature 1410°K.

If we add the further requirement that the δ-phase is unstable energetically to the β-phase at absolute zero then the maximum value of θ at which $H_0^{\delta} > H_0^{\beta}$ is $\theta = 380$. As we reduce θ below 380°K for the delta phase the electronic specific heat coefficient is also reduced such that at $\theta = 350$ the γ has been reduced to its allowable probable error deduced from heat content measurements of NAYLOR. We therefore place the limits of θ for the δ-phase between 350 and 380 and for the purpose of our calculations we choose $\theta = 370$ and $\gamma = 22{\cdot}5$. Table 9 lists the calculated thermodynamic functions for δ-manganese.

FREE ENERGY DIFFERENCES

Fig. 3 plots the calculated free energy differences of the various phases. Deviations from the

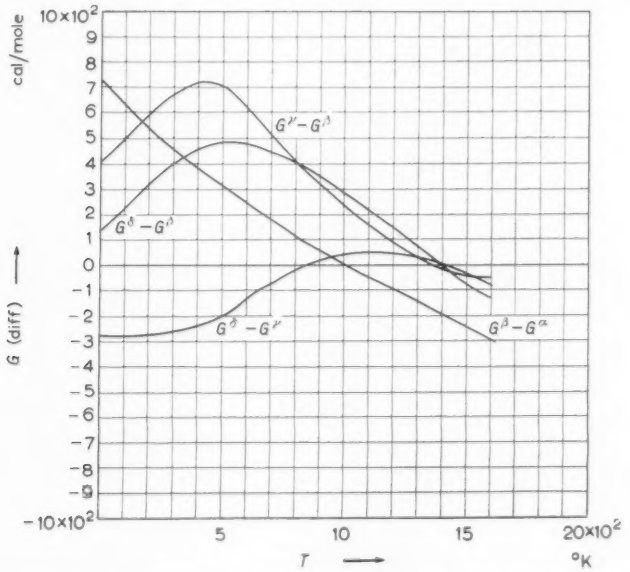


FIG. 3. The calculated free energy differences for the allotropic modifications of Mn. The energy of αMn at absolute zero is chosen as the zero energy base.

calculations of CHRISTIAN[15] occur but these are due to additional experimental data being available to the present authors.

DISCUSSION

It should be noted (Fig. 3) that if it were not for the presence of the complicated α- and β-phases, the body centred δ-phase would be the stable phase again at low temperatures in a similar fashion to iron. However recent views on the electronic structure of the body centred metals suggest that manganese would have one electron in a doubly degenerate *d* orbital. This would be subject to a Jahn–Teller distortion and the low temperature form of manganese would be body centred tetragonal. It is interesting that the complicated α phase consists of 59 per cent of the atoms having 1·54 μ_B and 41 per cent having zero μ_B thus tending to avoid atoms with a doubly degenerate *d* orbital. It is suggested that this is the key to the presence of the complicated manganese phases, i.e. the gain in intra-atomic exchange energy by pairing off the doubly degenerate *d* orbitals more than compensates for the distortional energy necessary to go from a simple structure to a complicated one. This is evidenced by the higher Debye temperatures for the complex structures but the lower crystalline energy at absolute zero.

Having separated out the free energy components we can now see which of the components are primarily responsible for the stability ranges of the various phases. αMn is most stable at low temperatures due to the low energy at absolute zero gained as we have said from the intra-atomic exchange energy due to pairing of *d* orbitals. The stability of the β-phase comes from the lower Debye θ of this phase as compared to the α-phase. It is further suggested that orbital pairing is not as complete in this phase as in the α-phase and that experiments (neutron diffraction on single crystals or paramagnetic diffuse scattering) will show moments closer to one Bohr magneton. The γ-phase (basically f.c.c. is not expected to show the high degree of hybridization and it is suggested that there are approximately 6·5 3*d* like electrons with 4·4 having spin up and 2·1 having spin down per atom. Its stability is a result of its magnetic free energy and low Debye temperature. Finally, the high temperature body centred δ-phase becomes stable relative to the γ-phase due to its high electronic specific heat. Its stability relative to the β-phase is a result of its low Debye temperature. Its low Debye temperature relative to its body centred cubic neighbors (370 compared to Cr 485 and Fe 430) is consistent with a Jahn–Teller distortion causing large vibrational amplitudes in the cubic phase. We should also like to point out that the electronic specific heat is approximately the same for the α-, β- and δ-phases. We lastly point out that this thermodynamic analysis indicates that in the absence of the γ-phase, the β- would transform to δ- at about 1395°K consistent with recent work of CHRISTIAN on alloying elements which close the γ-loop such as Cr and V. An extrapolation of the β–δ phase boundary in these alloys to pure Mn gives a value of ~ 1392 for the β–γ transition temperature.

Acknowledgements—We wish to express our appreciation for advice on these matters to DRS. J. CHRISTIAN, L. E. ORGEL, J. S. GRIFFITHS, H. BROOKS and J. S. LEACH.

APPENDIX

Derivation of the magnetic free energy expressions

In deriving these expressions we rely on the results obtained in reference 2 on many magnetic substances and the good agreement of these results with values of J/kT_c calculated from various statistical models. We find that the total energy of magnetization, *W*, divided by RT_c is a monotonic slowly varying function of spin and relatively independent of crystal structure. We empirically find

$$W = \int_0^\infty C_p(\text{mag})\, dT \cong RT_c\bar{s}^{\frac{1}{2}} \qquad (1a)$$

where $\bar{s}$ is the average magnitude of spin on the atoms. Furthermore we require the entropy to be given as

$$S \cong \int_0^\infty C_p(\text{mag})\, dT/T \cong R\ln(2\bar{s}+1). \qquad (2a)$$

Since we are only concerned with free energy or areas under the specific heat curve the exact shape of the specific heat curve is not critical and we merely require an approximately correct shape. We have chosen

$$C_p = AT^3 \quad \text{from } 0 \leqslant T \leqslant T_c \qquad (3a)$$

$$C_p = B/T^2 \quad \text{from } T_c < T \qquad (4a)$$

giving the correct spin wave dependence at low temperatures for antiferromagnets and the correct classical

approach to high temperatures for both ferro- and antiferromagnets. The fact that ferro and ferri magnets behave as $T^{3/2}$ in the low temperature spin wave approximation will not be critical since we do not intend to apply this expression at low temperatures. The above specific heat expressions exhibit a discontinuity at the critical temperature similar to that calculated in the cluster theories of order disorder but here again areas are our prime consideration. Lastly we require that the fractions of entropy and energy evolved at the Curie points shall agree with those experimentally observed.[13] Substituting 3A and 4A in 1A and 2A and eliminating A and B we obtain the free energy expressions given in the text. Fortunately these do give good agreement with the fractions as reported in reference 13.

REFERENCES

1. Weiss R. J. and Tauer K. J. *Phys. Rev.* **102,** 1490 (1956).
2. Hofmann J. A., Paskin A., Tauer K. J. and Weiss R. J. *J. Phys. Chem. Solids* **1,** 45 (1956).
3. Dunitz J. D. and Orgel L. E. To be published, *J. Phys. Chem. Solids* **3,** 20 (1957); Griffith J. S., *J. Inorg. Nucl. Chem.* **3,** 15 (1956).
4. Booth, Hoare and Murphy *Proc. Phys. Soc.* **68,** 830 (1955).
5. Shomate C. H. *J. Chem. Phys.* **13,** 326 (1945).
6. Armstrong L. D. and Grayson Smith H. *Canad. J. Res.* A**28,** 51 (1949).
7. Kasper J. S. and Roberts B. W. *Phys. Rev.* **101,** 537 (1956).
8. Naylor B. F. *J. Chem. Phys.* **13,** 329 (1945).
9. Bacon G. Private communication.
10. Smith L. P. Private communication.
11. Tauer K. J. and Weiss R. J. To be published *J. Phys. Chem. Solids* **4,** 135 (1957).
12. Yin Yuan Li *J. Chem. Solids* **17,** 447 (1949).
13. Paskin A. *J. Phys. Chem. Solids* **2,** 232 (1957).
14. Tauer K. J. and Weiss R. J. *Phys. Rev.* **100,** 1223 (1955).
15. Christian J. W. *J, Inst. Met.* **78,** 195 (1950).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 144–147.

ABSORPTION SPECTRA OF SOLIDS AND CHEMICAL BONDING—I

ARSENIC, ANTIMONY AND BISMUTH TRIOXIDES

W. P. DOYLE

Chemical Laboratories, Imperial College, London and University College, Dublin

(*Received* 13 *May* 1957)

Abstract—Measurements of the optical absorption of thin layers of arsenic, antimony and bismuth trioxides are reported. Each material is characterized by strong absorption in the ultra-violet; the threshold photon energies are respectively 4·0, 4·2, 3·2 eV and the absorption coefficient at higher energies is at least 10^5 cm^{-1}. The results are discussed qualitatively in relation to chemical bonding and related to the diffuse reflection spectra of the substances. In antimony trioxide there is some evidence for the production of excitons.

VOL. 4 1958

1. INTRODUCTION

THE optical absorption spectra of solids have been little discussed in their relation to chemical bonding. It is recognized that in the alkali halides although the ultra-violet bands are to be ascribed to absorption by the halogen ion the spectrum is considerably influenced by the particular cation[1] and it has been pointed out that whereas in the highly ionic alkali halides the absorption bands are very sharp in less ionic compounds the maxima are much more diffuse.[2] The complexity of the solid state makes quantitative interpretation of absorption data very difficult and only for the alkali halides have the absorption measurements been successfully interpreted in detail. The aim of the present work is to determine the absorption spectra of a relatively large number of substances and by qualitative comparison and correlation to attempt to relate absorption data to type of chemical bonding.

The course of optical absorption in thin films of solids has been determined only for the following compounds (to eliminate the complication of intra-atomic electronic transitions, only compounds of non-transition elements are considered); alkali halides,[3,4] hydrides,[5] and oxides;[6] the halides of monovalent silver and copper;[7] magnesium fluoride, chloride and bromide, calcium fluoride and chloride,[4] strontium oxide,[8] barium chloride,[4] bromide,[9] and oxide;[10] zinc fluoride,[11] iodide,[9] oxide and sulphide,[12] cadmium chloride, bromide, iodide,[9] oxide, and sulphide;[12] monovalent indium chloride;[9] lead chloride, bromide, iodide,[9] sulphide, selenide, and telluride;[13] arsenic trisulphide.[14] In view of the paucity of existing data the initial choice of substances is somewhat arbitrary. Optical absorption in solids appears to be determined primarily by the anions and the common simple anions are those of Groups VI and VII of the periodic table. Most existing data refer to substances containing Group VII, halide, anions and so it was decided to extend the data on substances containing Group VI anions, principally oxides and sulphides. Here absorption measurements on the oxides of arsenic, antimony and bismuth are presented.

2. EXPERIMENTAL AND RESULTS

Films sufficiently thin for the determination of optical absorption in solids are usually prepared by vacuum evaporation. The evaporation chamber used consisted of a brass cylinder closed at either end by circular brass plates in which were grooves holding O-ring vacuum seals. The upper plate carried a fused quartz disc as support for the evaporated layer and the lower plate the leads for the heating element. The latter was a small boat of 0·002 in. thick tungsten strip and was cleaned before use by heating in hydrogen. Evaporation was carried

out at a pressure of approximately 10^{-5} mm of mercury.

In vacuo both arsenic and antimony trioxides sublime without melting and for such substances it is necessary to use fused lumps of material.[15] Arsenic trioxide, 99·8 per cent pure, was heated to 400°C in a sealed tube and on cooling a compact mass of the vitreous form was obtained.[16] On evaporation uniform, turbid, white films were obtained; the reason for the turbidity of these films is not known. Merck "pure" antimony trioxide was prefused by rapid heating in a nitrogen atmosphere. On evaporation clear colourless films were obtained. Attempts to prepare films of bismuth trioxide by direct evaporation resulted in the deposition of a film of bismuth metal due to decomposition of the oxide under the evaporation conditions. Bismuth trioxide films were prepared by vacuum evaporation of bismuth, 99·9 per cent pure, and these metallic films when heated to 100°C in a stream of oxygen were rapidly converted to turbid, white films of the trioxide. Turbidity of films prepared in this way might be expected as the expansion occurring when the metal is converted to oxide probably breaks up the film into minute particles which cause light scattering.

The three substances examined are chemically stable in air and therefore it was not necessary to determine the spectrum *in vacuo* or in a protective atmosphere. The absorption of a film was determined in a Beckmann Model DU spectrophotometer by measuring the optical density (the decadic logarithm of the ratio of the intensity of the incident light, I_0, to that of the transmitted light, I_t,) of the specimen and quartz disc mount relative to that of a blank quartz disc. The blank disc and the film mount had practically identical absorption in the spectral range studied. In measuring the optical density of a film no corrections were made for reflection or scattering losses. For the determination of the shape and position of the absorption bands this is unimportant. The only disadvantage occurs in the long-wave branch of the absorption spectrum where absorption is simulated due to reflection and scattering losses. In this region of low absorption very thin films had practically zero optical density showing that reflection losses are negligible and that the apparent absorption observed in the thicker films is due almost entirely to scattering. The wavelength dial of the spectrophotometer was calibrated against a mercury arc, and in addition, an absorption curve was run on potassium iodide. The position of the first absorption peak and the shape of the curve for potassium iodide agreed closely with the results reported by HILSCH and POHL.[3] In the case of each trioxide, the absorption of at least three films covering a variation in thickness of a factor of at least two was measured and the absorption curves were found to be reproducible.

The thickness of a film was determined approximately by weighing and assuming the density of a film to be the same as that of the bulk material. All three trioxides exist in at least two polymorphic forms but, except for antimony trioxide, the forms do not differ greatly in density so that in calculating the thickness the crystal form is not very important. Antimony trioxide when sublimed exists as the cubic polymorph[17] and so the appropriate density could be used in calculating the thickness. From the thickness, t, of a film absolute values of the absorption coefficient, k, (defined by the equation $I_t = I_0 \cdot e^{-kt}$) were calculated. It is unlikely that the assumption of bulk density for a film introduces significant error because it might be expected that, to a first approximation, the absorption coefficient would be proportional to the density if the latter differed from the ideal bulk value. Thus when the thickness of a film is determined by weighing, errors due to the assumption of bulk density tend to cancel out.

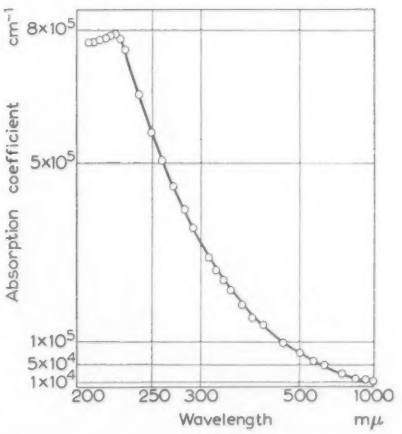


FIG. 1. Absorption spectrum of arsenic trioxide film, 300 Å thick.

The absorption spectra, from 205mμ to 1000mμ, of typical films of the three trioxides are shown in the figures and the thickness of each film is indicated. For arsenic trioxide the light scattering of the films masks the onset of strong absorption.

Extrapolation of the steep portion of the curve indicates that the absorption coefficient will reach zero at about 310 mμ, and this may be taken as the threshold of strong absorption which then rises

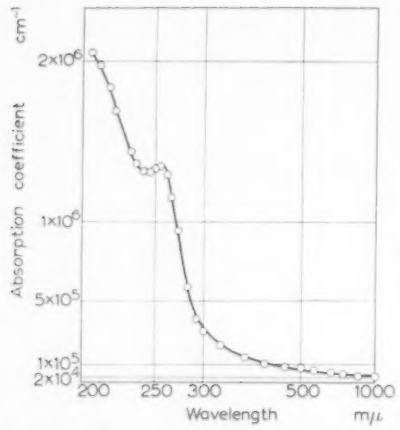


FIG. 2. Absorption spectra of antimony trioxide film, 60 Å thick.

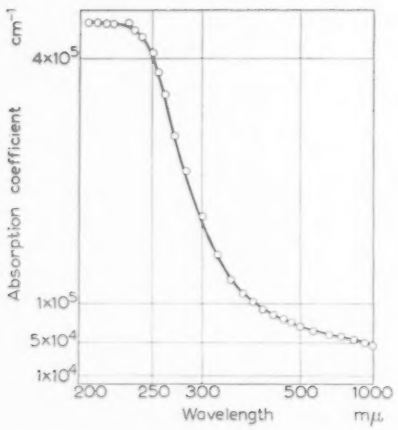


FIG. 3. Absorption spectrum of bismuth trioxide film, 400 Å thick.

to a maximum at 225 mμ. The main features of the spectrum of antimony trioxide are the sharp threshold at 295 mμ and the well-defined step from 260 to 240 mμ. In bismuth trioxide, the absorption increases rapidly from about 390 mμ and rises to a maximum at 230 mμ. The course of the absorption at wavelengths shorter than 230 mμ was not completely reproducible. For some films it remained constant from 230 to 205 mμ and for others it decreased slightly in this region. This slight variation is presumably associated with differences in the conditions of conversion of the metallic films to the trioxide.

3. DISCUSSION

The absorption spectra of insulators are usually characterized by very strong absorption at short wavelengths with a sharp edge on the long wavelength side. The absorption coefficient is usually of the order of 10^5–10^6 cm^{-1}. The spectra of the three trioxides are also of this general type. For all three substances the magnitude of the absorption coefficient in the regions of strong absorption shows that absorption must be attributed to the bulk lattice and not to defects or impurities. As for other insulators, lattice absorption almost certainly corresponds to electronic transitions between the highest filled energy band and the lowest empty band. To a rough approximation this corresponds to the transfer of an electron from O ions to As, Sb or Bi ions. To confirm the nature of the transition involved the energy for the transfer of an electron from the anion to the cation should be calculated, as done for the alkali halides.[18] However, this calculation involves the Madelung constant, which for the crystal lattices concerned has not been computed.

As indicated above, the absorption data at present available are insufficient for detailed correlation of spectra and chemical bonding and only a brief analysis is given. The interpretation of the present data is based on the plausible assumption that the greater the ionic character of a chemical bond the more closely are the valency electrons of the cation associated with the anion. Thus, if the absorption process consists of an electron transfer from anion to cation, then, neglecting factors such as crystal structure, it should follow that the more ionic the bond the greater the energy of the transition. Hence the absorption spectra of ionic compounds may be expected to lie at shorter wavelengths than those of more covalent compounds. This is supported by comparison of the spectra of arsenic and antimony oxides with those of the corresponding trisulphides. In the oxides, as might be expected from the greater electronegativity of oxygen relative to that of sulphur, the threshold

energy for lattice absorption is at shorter wavelengths than in the corresponding sulphides. As_2O_3 310 mμ, As_2S_3 500 mμ;[14] Sb_2O_3 295 mμ, Sb_2S_3 525 mμ.[19] The energy difference for both pairs of substances is appreciable, about 1·6 eV, corresponding to the appreciable difference in electronegativity, 1·0, between oxygen and sulphur.[20] Another difference between the oxide and sulphide spectra is that the absorption curve of the oxide is much steeper than that of the corresponding sulphide which also may be associated with the much greater ionic character of oxides.[2]

The step in the spectrum of antimony trioxide is similar to a step occurring in the spectrum of barium oxide in the wavelength region 290–250 mμ.[10] ZOLLWEG has resolved the step in barium oxide into two peaks by measuring the spectra at liquid nitrogen temperatures and by employing greater resolution, and has suggested that they were caused by exciton absorption.[21] It seems probable that the step in the antimony oxide spectrum may be attributed to the same cause, although no explanation can be given why similar steps do not occur with arsenic and bismuth oxides. In this connection it is noteworthy that in the spectrum of strontium oxide[8] there is no step analogous to that in barium oxide where ZOLLWEG's work provides strong evidence for the exciton interpretation. It is very improbable that such a step is associated with differences in lattice structure because strontium and barium oxides have the same, rock-salt, lattice. The possibility that the antimony trioxide step is associated with antimony atoms or ions appears unlikely because no such step is observed in the spectrum of antimony trisulphide.[19]

For antimony trioxide, the threshold of strong absorption is confirmed by the known diffuse reflection threshold; it is generally agreed that a rapid decrease in the diffuse reflection coefficient corresponds to the beginning of strong absorption as determined by transmission measurements on thin films. The diffuse reflection threshold of rhombic antimony trioxide is 373 mμ and that of the cubic form is 300 mμ.[22] Films of antimony trioxide have the cubic structure and the absorption threshold is at 295 mμ in close agreement with the diffuse reflection threshold of the cubic form. For bismuth trioxide, the diffuse reflection threshold of the rhombic form is 436 mμ and of the cubic form, 387 mμ.[22] The absorption threshold of the thin films prepared as described is 390 mμ suggesting that they possess the cubic form.

Acknowledgements—The Author is greatly indebted to Dr. J. A. KITCHENER of Imperial College, London for much valuable assistance and to Prof. T. S. WHEELER of University College Dublin for his continued encouragement.

REFERENCES

1. HILSCH R. and POHL R. W. *Z. Phys.* **48,** 384 (1928).
2. FESEFELDT H. *Z. Phys.* **64,** 741 (1930).
3. HILSCH R. and POHL R. W. *Z. Phys.* **59,** 812 (1929).
4. SCHNEIDER E. G. and O'BRYAN H. M. *Phys. Rev.* **51,** 293 (1937).
5. RAUCH W. *Z. Phys.* **111,** 650 (1938–9).
6. RAUCH W. *Z. Phys.* **116,** 652 (1940).
7. FESEFELDT H. and GYULAI Z. *Nachr. Ges. Wiss. Göttingen* 226 (1929).
8. SPROULL R. L. *Phys. Rev.* **78,** 630 (1950).
9. FESEFELDT H. *Z. Phys.* **64,** 741 (1930).
10. TYLER W. W. *Phys. Rev.* **76,** 1887 (1949).
11. PARKINSON W. W. and WILLIAMS F. E. *J. Chem. Phys.* **18,** 534 (1950).
12. MOLLWO E. *Reichsberichte für Physik* (formerly *Phys. Z.*) **1,** 1 (1944).
13. GIBSON A. F. *Proc. Phys. Soc. Lond.* B **63,** 756 (1950).
14. DOYLE W. P. *Proc. Phys. Soc. Lond.* B **69,** 865 (1956).
15. STRONG J. *Modern Physical Laboratory Practice.* Blackie, London (1938).
16. DEBRAY H. *C.R. Acad. Sci., Paris* **58,** 1209 (1864).
17. TERREIL H. *Ann. Chim.* (*Phys.*) **7,** 350 (1866).
18. VON HIPPEL A. *Z. Phys.* **101,** 68C (1936).
19. DOYLE W. P. To be published in *J. Phys. Chem. Solids.*
20. PAULING L. *The Nature of the Chemical Bond.* Cornell University Press, New York (1945).
21. ZOLLWEG R. J. *Phys. Rev.* **97,** 288 (1955).
22. BERTON A. *C.R. Acad. Sci., Paris* **208,** 1898 (1939).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 148–153.

INFRARED SPECTRA OF GROUP III ACCEPTORS IN SILICON

H. J. HROSTOWSKI and R. H. KAISER

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

(*Received* 5 *August* 1957)

Abstract—Infrared spectra of boron-, aluminum-, gallium-, and indium-doped silicon have been investigated at 4·2°K, using improved techniques. Many new absorptions were observed or resolved. The term schemes for excited states are now found to be consistent for all acceptors and to agree well with the term scheme calculated from the effective-mass theory. The discrepancies in the gallium spectra are shown to result from the presence of oxygen. The similarity of boron to a hydrogenic center is removed by observation of a previously unobserved low-energy transition.

INTRODUCTION

DISCRETE electronic transitions of a bound hole from the ground state of an acceptor atom to excited states were first observed in the low-temperature infrared spectra of boron-doped silicon by BURSTEIN and co-workers.[1] Further investigations by BURSTEIN *et al.*[2] and NEWMAN[3,4] contributed greatly to the experimental knowledge of the excited states of Group III acceptors.* However, these results also indicated marked differences in the spectra of different impurities. Simultaneous theoretical development progressed from the simple hydrogenic model† for an acceptor center to the more exact effective-mass theory[5–7] based on knowledge of the valence band of silicon obtained from cyclotron-resonance experiments.

The present effective-mass theory neglects the effects of the specific nature of the impurity atom in the proximity of the acceptor center. However, this approximation is expected to have a large influence only on the position of the ground state relative to the top of the valence band. Accordingly, the term schemes for excited states should be very similar for different acceptor atoms. Previous data did not support this expectation.

Improvements in the control of impurities in silicon and in experimental techniques have led us to re-investigate acceptor spectra. We have observed or resolved many transitions not reported previously. The results are in rather good agreement with the predictions of the effective-mass theory, but vary somewhat from acceptor to acceptor.

EXPERIMENTAL PROCEDURE

All measurements were taken at 4·2°K with a Perkin-Elmer Model-112 spectrometer in a manner described elsewhere.[8] NaCl, CsBr, and CsI optics were used. The CsI prism was calibrated as suggested by ACQUISTA and PLYLER.[9] Spectral resolution is indicated in each figure. Major emphasis was placed on the accurate location and resolution of the absorption bands. Errors in frequency measurement were minimized by frequent calibrations in the appropriate spectral regions. No attempt was made to determine absorption cross-sections, since small discrepancies in per cent transmission lead to large errors in absorption coefficients when absorption is very weak or very strong. Although the observed per cent transmission may be in error by as much as five units, all bands reported were experimentally reproducible.

Samples were taken from suitably doped single crystals prepared by floating-zone techniques or pulled from the melt. Approximate resistivities and net acceptor concentrations are given in each figure. In all cases the acceptor density was kept below

* Since we are dealing only with Group III acceptors, we shall in future refer only to acceptors.

† Attempts to interpret acceptor atoms as hydrogenic centers are described by BURSTEIN *et al.*[1,2]

about 10^{16}/cm^3 to avoid broadening effects arising from electronic interaction of the impurity centers.[4]

EXPERIMENTAL RESULTS

Figs. 1(a) to (d) show the spectra observed for boron-, aluminum-, gallium-, and indium-doped silicon, respectively, with per cent transmission plotted against frequency in cm^{-1}. The solid curves were obtained with floating-zone samples which contained no detectable oxygen.[10] The dashed curve in Fig. 1(c) also shows results for a pulled gallium-doped sample with an oxygen concentration greater than 10^{17} atoms/cm^3. The absorption bands of boron, aluminum, and indium showed no frequency shift within experimental error when the oxygen concentration was varied by a factor of at least 10^2. This is also true for all gallium bands except the 523·1 cm^{-1} band in "oxygen-free" silicon, which shifted to 517·3 cm^{-1} for silicon highly contaminated with oxygen. Since this variable frequency band is the only acceptor absorption which is dependent on oxygen concentration, we conclude that it is not associated with an electronic transition in the gallium spectra. The anomalous width of the 498·6 cm^{-1} gallium absorption and the origin of the 523·1 cm^{-1} band are discussed in the next section.

The spectra of all acceptors are strikingly similar. Only very minor differences are apparent. With the exception of the great width of the 498·6 cm^{-1} gallium absorption, the discrepancies are: (1) the larger number of observed transitions to more highly excited states in the boron spectrum, and (2) the intensity reversal in the two aluminum bands near 520 cm^{-1}. A one-to-one correspondence of bands in different spectra is easily made, in agreement with the theoretical expectation that the

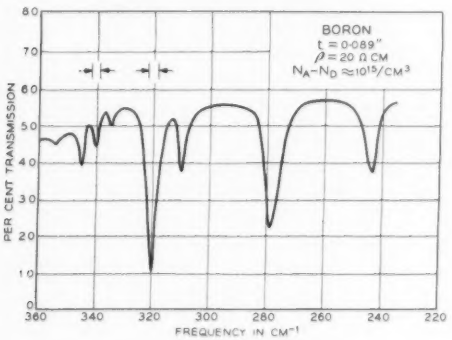


FIG. 1(a).

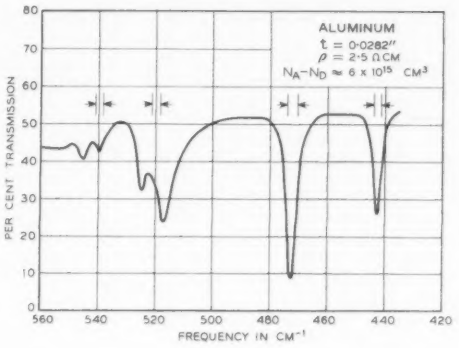


FIG. 1(b).

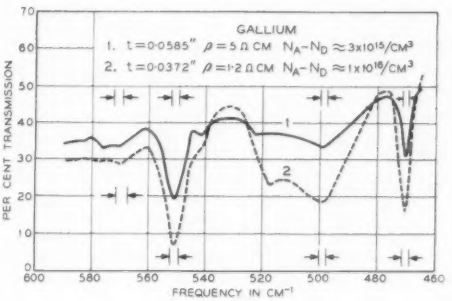


FIG. 1(c).

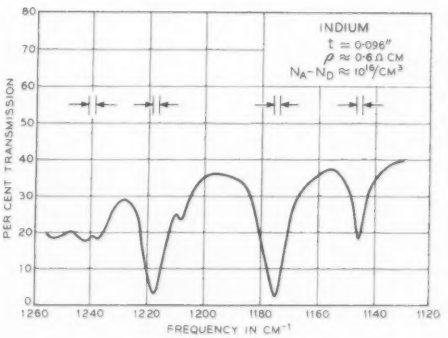


FIG. 1(d).

FIG. 1. The transmission spectra of (a) boron-, (b) aluminum-, (c) gallium-, and (d) indium-doped silicon at 4·2°K. All samples were prepared by floating-zone techniques except for sample 2 in (c), which was from a pulled crystal. Approximate spectral slit widths are given except for the CsI region (below 300 cm^{-1}), in which the spectral slit widths are probably comparable to the observed band widths.

Table 1. *Experimental acceptor spectra in units of* 10^{-3} *eV*

Transition	Boron	Aluminum	Gallium	Indium
$1 \rightarrow 2$	$30{\cdot}16 \pm 0{\cdot}12$	54·91	58·26	142·06
$1 \rightarrow 3$	$34{\cdot}52 \pm 0{\cdot}12$	58·57	61·80	145·73
$1 \rightarrow 4$	38·40	64·08	67·15	149·75
$1 \rightarrow 5$	39·65	64·99	68·26	150·88
$1 \rightarrow 6$	41·46	66·90	70·68	153·36
$1 \rightarrow 7$	42·13	67·55	$71{\cdot}39 \pm 0{\cdot}15$	153·96
$1 \rightarrow 8$	42·74	$68{\cdot}51 \pm 0{\cdot}12$	$72{\cdot}28 \pm 0{\cdot}15$	$155{\cdot}41 \pm 0{\cdot}2$
$1 \rightarrow 9$	43·85			

term schemes for excited states of all acceptors should be very similar.

The energies of the observed transitions (in eV) are listed in Table 1 with the variable energy absorption of the gallium spectra excluded. Less than half of these transitions were observed or resolved in previous investigations. Except where otherwise noted, the average experimental error is 3×10^{-5} eV. The error in the first two boron bands arises from uncertainty in the calibration of the CsI prism.

We observed one more transition for boron than for aluminum, gallium, and indium, since the higher boron transitions occur in a spectral region of high resolution. More careful work employing gratings may resolve more transitions for the other acceptors and perhaps even higher-energy boron transitions.

The levels involved in these transitions are not the simple *s*, *p*, *d*, etc., states of isolated atoms, but complicated mixtures of such states.[11] Although the four lowest excited states become hydrogenic $2p$ functions for an isotropic effective mass,[11] the nature of the higher levels is still unknown. For this reason we make no attempt to designate the transitions by the nature of the states involved. The states are labelled in numerical order, increasing from the ground state to the continuum. We assume that all transitions observed originate from the ground state.

THE EFFECT OF OXYGEN

It has been shown that silicon crystals pulled from fused-quartz crucibles contain large concentrations of oxygen.[8,10] The ratio of oxygen atoms to impurity centers was about 10^2 in samples previously used to investigate acceptor spectra in the low concentration limit. Oxygen contamination masks the first transition in the indium spectrum. A very intense silicon–oxygen stretching motion absorbs in this region.[10] At 4·2°K this vibration occurs at 1135·5 cm^{-1} and has a maximum absorption coefficient of about 20 cm^{-1} for pulled rotated crystals.[8] Therefore previous experiments with pulled silicon crystals have failed to locate the lowest indium transition, which is readily detected in floating-zone ("oxygen-free") silicon. Verification that the 1146·1 cm^{-1} absorption in Fig. 1(d) is an indium transition is obtained by the similarity of its intensity behavior on cooling to that of the corresponding absorptions of the other acceptors. Furthermore, the low-temperature fine structure associated with the silicon–oxygen stretching motion is absent.

The variable frequency absorption in the gallium spectra suggests an explanation for the anomalous width of the gallium band at 498·6 cm^{-1}. In oxygen-contaminated silicon (10^{17} atoms/cm^3) with acceptor density too low to detect by absorption measurements, a weak silicon-oxygen bending vibration occurs at 517·3 cm^{-1} at 4·2°K.[8] This band has a much lower intensity than the band observed at this exact frequency in pulled gallium-doped crystals in which the gallium concentration is only 1 per cent of the oxygen concentration. In spite of the difference in intensity, the frequency coincidence indicates that the 517·3 cm^{-1} band in the gallium spectrum is this silicon–oxygen absorption. For comparable concentrations of gallium and oxygen, the 517·3 cm^{-1} band shifts to 523·1 cm^{-1}. This shift may indicate that gallium is directly bonded to oxygen in Si–O–Ga units. The broadening of the 498·6 cm^{-1} gallium frequency apparently results from an interaction of the

VOL. 4 1958

silicon–oxygen bending mode with the electronic transition of a bound hole from the ground state to the second excited state of gallium. The intensity of the silicon–oxygen vibration is greatly enhanced by the interaction.* These results lead us to believe that the variable frequency band is therefore not associated with the true gallium spectrum. Although similar broadening might be expected in the aluminum bands at 517·0 and 524·3 cm^{-1}, none is observed. The reason for the absence of broadening here is not clear. However, the electronic nature of the excited states involved probably forbids interaction with the silicon–oxygen vibration. While the second excited gallium state is four-fold degenerate, the third and fourth excited aluminum states are both two-fold degenerate.[11]

ENERGY-LEVEL SPACINGS

Table 2 lists the acceptor-level spacings and the average values for the four impurities. The experimental errors are included. These values give the positions of all excited states relative to the first (lowest lying) excited state. The data show several important features:

(*a*) The most remarkable observation is the overall agreement for different acceptors. The spacings of corresponding excited levels are very similar. When experimental errors are considered, the largest discrepancies amount to about 20 per cent of the average values. With one exception, the spacings of corresponding energy levels of aluminum and gallium are almost identical within the errors of measurement. This result is not unexpected, since both elements have a covalent tetrahedral radius of 1·26 Å,† which is close to that of silicon (1·17 Å). The differences between the highest and lowest excited states range from $13{\cdot}35\times10^{-3}$ eV for indium to $14{\cdot}02\times10^{-3}$ eV for gallium. The maximum deviation from the average value is less than 3 per cent. We note also that indium, which has a comparatively high ionization energy, agrees reasonably well with the other acceptors.

(*b*) The differences between acceptors are emphasized by comparison of the term schemes shown in Fig. 2. Here the ground state is ignored and the lowest-energy excited state is assumed to lie a constant (unknown) energy above the top of the valence band. This method of plotting is used for two reasons: (1) the agreement between the lower excited states of aluminum and gallium is emphasized, and (2) there is ample theoretical justification for assuming a constant energy difference between any excited state and the continuum.[13] The energy-level spacings for boron and indium are reasonably different from those of aluminum and gallium, which also differ to a lesser extent.

(*c*) Another important feature is shown by the average spacings given in Table 2. Instead of decreasing regularly, the spacings between adjacent levels alternately increase and decrease. This pattern is followed up to the highest separation observed for all four acceptors.

* The intensity increase is so great that the 517·3 cm^{-1} band on the broadening of the 498·6 cm^{-1} gallium band in gallium-doped silicon may prove to be more sensitive indications of oxygen contamination than the 1135·5 cm^{-1} silicon–oxygen band.

† Values of the tetrahedral radii are from PAULING.[12]

Table 2. Observed energy-level spacings in units of 10^{-3} *eV*

ΔE	Average	Boron	Aluminum	Gallium	Indium
3—2	3·81	4·36±0·24	3·66±0·04	3·54±0·06	3·67±0·10
4—3	4·69	3·88±0·13	5·51±0·06	5·35±0·08	4·02±0·10
5—4	1·10	1·25±0·02	0·91±0·06	1·11±0·08	1·13±0·11
6—5	2·16	1·81±0·02	1·91±0·06	2·42±0·08	2·48±0·11
7—6	0·66	0·67±0·02	0·65±0·08	0·71±0·19	0·60±0·10
8—7	0·98	0·61±0·04	0·96±0·16	0·89±0·30	1·45±0·26
9—8	—	1·11±0·07			
8—2			13·61±0·14	14·02±0·17	13·35±0·24
	13·68				
9—2		13·69±0·16			

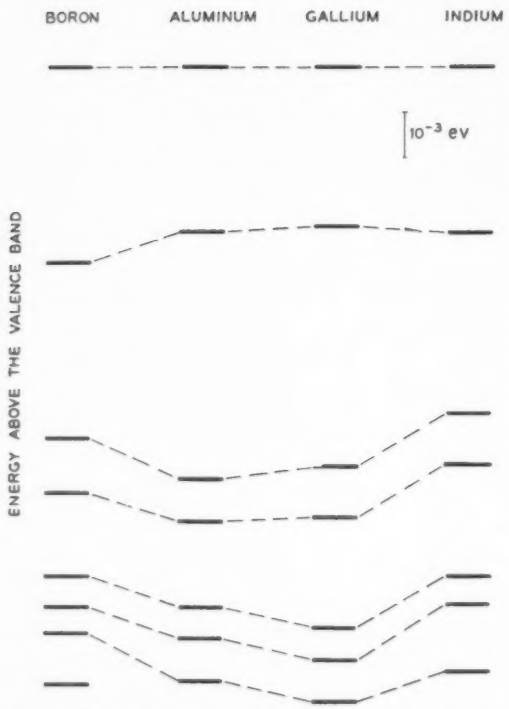


FIG. 2. The term schemes for excited states of boron, aluminum, gallium, and indium. The lowest excited states are assumed to be a constant (unknown) energy above the top of the valence band.

COMPARISON WITH THEORY

Although the simple hydrogenic model for an impurity center has been superseded by the effective-mass theory, based on experimental knowledge of the band structure, previous experiments indicated that boron could be interpreted as a hydrogen-like center. Three bands observed at 0·034, 0·040, and 0·043 eV agreed remarkably well with those calculated from the hydrogenic model, using a reasonable value for the effective hole mass.[1] Extension of the boron spectra to lower energy now shows the presence of another band at 0·03016 eV, (as shown in Fig. 1(a).) This observation proves the similarity of boron to other acceptors which could not be described by the simple hydrogenic model. Thus none of the data can be interpreted by this model, which is at best only a rough approximation to the true nature of impurity centers.

According to the effective-mass theory, the excited acceptor states are determined by a set of six coupled differential equations. The energy of any particular level depends on the shape of the constant-energy surfaces in the valence band and the spin-orbit splitting.[5,6] Although the spin-orbit splitting has not been experimentally determined theoretical estimates give a value comparable to the ionization energies of acceptors. The constants which determine the constant-energy surfaces are not very well known. One of these, B, is even of uncertain sign. In spite of these difficulties, KOHN and SCHECHTER[11] have calculated an energy-level diagram for the lowest five states of an acceptor. An important result of this treatment is that the degeneracy of these states is lower than expected from the hydrogenic model. They are either two-fold or four-fold degenerate. Thus the observed spectra should contain more bands than that of a hydrogen-like center. The experimental data (Figs. 1(a) to (d)) bear out this conclusion.

Table 3. Calculated and observed energy-level spacings in units of 10^{-3} *eV*

ΔE	Calculated		Observed*
	B > O	B < O	
3 – 2	4·3	4·0	3·81
4 – 3	2·9	2·8	4·69
5 – 4	2·0	1·5	1·10

* Average values.

The calculated separations between the four lowest excited levels for both signs of B are given in Table 3, along with the average experimental values. In view of the uncertainties involved in the calculations, the agreement is quite good for two of the three spacings. The experimental results favor a negative value of B. It is clear that the effective-mass theory gives a much improved model for excited acceptor states. Even indium, with a thermal-ionization energy of 0·16 eV,[14] is in reasonably good agreement with theory. However, quantitative agreement cannot be expected until the specific effects of the acceptor atoms or other presently unknown factors are considered.

IONIZATION ENERGIES

Without exact knowledge for each acceptor of the position of one excited state relative to the continuum, the optical ionization energy, E_0, cannot be determined from measurements of this type. However, it is unlikely that the highest excited state is much more than 10^{-3} eV above the top of the valence band. Thus the position of this level sets a reasonably close lower limit to the true E_0. These approximate values of E_0 are given in Table 4. We include previous values of E_0 obtained by BURSTEIN *et al.*[2] assuming that the "4*p* hydrogenic" level lies 3×10^{-3} eV above the top of the valence band. The thermal-ionization energies, E_T, obtained from the temperature variation of the Hall coefficient[14] are also listed in Table 4.

Table 4. "Optical" and thermal ionization energies in units of 10^{-3} *eV*

Element	E_0	E_0[2]	E_T[14]
Boron	43·85	46	45
Aluminum	68·51	67	57
Gallium	72·28	71	65
Indium	155·41	154	160

For boron and indium, optical and thermal results agree within experimental error. The values for aluminum and gallium differ considerably. It has been suggested that the differences result from the presence of uncompensated boron centers which would give thermal values lower than the true values.[2] The magnitude of the discrepancies is such that almost equal concentrations of aluminum (or gallium) and uncompensated boron must be present to account for the thermal values. The thermal measurements, however, were taken on samples with relatively small concentrations of uncompensated boron present.[15] There is then a real difference between E_0 and E_T for aluminum and gallium. It is interesting to note that $E_0 > E_T$ for the elements which give the best agreement with the effective-mass theory. For boron and indium, which give poorer agreement with theory, $E_0 \approx E_T$. The source of these discrepancies is as yet unknown.

Acknowledgements—We have profited much from many discussions with W. KOHN. We are particularly grateful to H. E. BRIDGERS, Jr., E. KOLB, and H. C. THEUERER for the preparation of the crystals used in this work.

REFERENCES

1. BURSTEIN E., BELL E. E., DAVISSON J. W. and LAX M. *J. Phys. Chem.* **57,** 849 (1953).
2. BURSTEIN E., PICUS G., HENVIS B. and WALLIS R. *J. Phys. Chem. Solids* **1,** 65 (1956).
3. NEWMAN R. *Phys. Rev.* **99,** 465 (1955).
4. NEWMAN R. *Phys. Rev.* **103,** 103 (1956).
5. KITTEL C. and MITCHELL A. H. *Phys. Rev.* **96,** 1488 (1954).
6. LUTTINGER J. M. and KOHN W. *Phys. Rev.* **97,** 869 (1955).
7. KOHN W. and SCHECHTER D. *Phys. Rev.* **99,** 1903 (1955).
8. HROSTOWSKI H. J. and KAISER R. H. *Phys. Rev.* **107,** 966 (1957).
9. ACQUISTA N. and PLYLER E. K. *J. Opt. Soc. Amer.* **43,** 977 (1953).
10. KAISER W., KECK P. H. and LANGE C. F. *Phys. Rev.* **101,** 1264 (1956).
11. SCHECHTER D. forthcoming Ph.D. thesis, Carnegie Institute of Technology, Pittsburgh, Pennsylvania (Private communication from W. KOHN).
12. PAULING L. *The Nature of the Chemical Bond.* Cornell University Press, Ithaca, New York (1954),
13. KOHN W. Private communication.
14. MORIN F. J., MAITA J. P., SHULMAN R. G. and HANNAY N. B. *Phys. Rev.* **96,** 833 (1953).
15. MORIN F. J. Private communication.

New semiconductors

(*Received* 22 *July* 1957)

SEMICONDUCTORS are characterized by a large covalent-bond component. Intermetallic compounds that are semiconductors contain at least one element from Group 4, 5, or 6 of the Periodic Table. Such elements tend to form covalent bonds. Our synthesis of new semiconductors[1] has been based on substituting the stoichiometric amounts of chemically similar elements for those present in minerals containing at least one of the Group 4, 5, or 6 elements. The compounds with the ideal formulae of the minerals have also been synthesized. This report covers the synthesis of the semiconducting compounds $CuPbAsS_3$ (seligmannite[2]), $AgAsSe_2$, and $AgAsTe_2$. A discussion of the structures of $AgSbSe_2$ and $AgSbTe_2$ is also given, together with some qualitative doping data on the latter compounds.

Experimental

The compounds were prepared by melting the proper amounts* of the component elements in sealed quartz tubes containing a nitrogen pressure of approximately two-thirds of an atmosphere. Heating was done slowly so that the major portion of the alloying was carried out at temperatures at which the vapor pressures of arsenic, selenium, and sulfur were low. The average weight of the ingot made was 50 g.

Table 1. *Melting-point data*

$CuPbAsS_3$	460°C
$AgAsSe_2$	390°C
$AgAsTe_2$	325°C

Microscopic examination and thermal analysis showed that the compounds $CuPbAsS_3$, $AgAsSe_2$, and $AgAsTe_2$ were single-phase. X-ray powder photographs established that $CuPbAsS_3$ had been synthesized. The structures of $AgAsSe_2$ and $AgAsTe_2$ have not, as yet, been determined. The melting points of the three compounds are listed in Table 1.

Preliminary measurements of resistivity versus temperature on impure polycrystalline specimens indicate that the intrinsic energy gaps of the above compounds lie in the range 0·8–1·0 eV.

Some structural and doping results on $AgSbSe_2$ *and* $AgSbTe_2$

The ternary compounds $AgSbSe_2$ and $AgSbTe_2$† have the NaCl-type structure with two formula units per unit cell. The quantitative X-ray intensity data indicate conclusively that the structure is only partially disordered, with silver and antimony atoms indistinguishable in the positions (000, $0\frac{1}{2}\frac{1}{2}$, $\frac{1}{2}0\frac{1}{2}$, $\frac{1}{2}\frac{1}{2}0$) of the Na^+ ions of the NaCl structure, and the selenium or tellurium atoms at the positions ($\frac{1}{2}\frac{1}{2}\frac{1}{2}$, $\frac{1}{2}00$, $0\frac{1}{2}0$, $00\frac{1}{2}$) of the Cl^- ions. The lattice constants and X-ray and measured densities (by liquid-displacement technique) are given in Table 2. Annealing experiments at various temperatures and for various periods of time did not yield any other phases.

Table 2. *Lattice constants and density data*

Compound	Lattice constant (Å)	Density (g/cm³)	
		Measured	X-ray
$AgSbSe_2$	5·786±0·003	6·69	6·60
$AgSbTe_2$	6·078±0·003	7·12	7·12

Excess antimony, zinc, cadmium, or manganese changes ordinarily *p*-type $AgSbSe_2$ to *n*-type. Excess silver increases *p*-type conductivity. It is worth noting that, from a knowledge of the structure, one could predict the result of excess

* To anticipate vaporization losses, usually 0·5 g arsenic, selenium or sulphur was added in excess of the stoichiometric amounts. However, in nearly all cases, practically no weight losses were noted when the ingots were weighed.

† These compounds are isostructural with the high-temperature form of $AgBiS_2$[3,4] and $AgBiSe_2$. We have synthesized both of these compounds, the latter being new. The low-temperature forms of both are isostructural.

antimony or silver. It is likely that zinc, cadmium, and manganese would also prefer positions occupied by silver and antimony atoms.

Mobility measurements

We have begun a program of zone refining the compounds with structures related to that of NaCl. We have obtained the very promising result of 1500 cm^2/V sec for the hole mobility of a zone-refined single crystal of $AgSbSe_2$ at 23°C. Further measurements of fundamental quantities are under way.

A paper describing in detail the structures of the compounds related to the NaCl type is in preparation.

Bell Telephone Laboratories Inc.
Murray Hill, New Jersey
J. H. WERNICK
S. GELLER
K. E. BENSON

VOL. 4 1958

REFERENCES

1. WERNICK J. H. and BENSON K. E. *J. Phys. Chem. Solids* To be published.
2. HELLNER E. and LEINEWEBER G. *Z. Kristallogr.* **107,** 150 (1956).
3. RAMDOHR P. *S.B. preuss. Akad. Wiss., Phys.-math. Kl.* **6,** 71 (1938).
4. GRAHAM A. R. *Amer. Min.* **36,** 436 (1951).

Extension of infrared spectra of III–V compounds by lithium diffusion

(*Received* 29 *August* 1957)

PRESENT techniques of preparing III–V semiconducting compounds often yield materials with relatively high impurity concentrations. Free-charge-carrier absorption is often so strong that lattice vibrations cannot be observed by transmission measurements. As a result many of the optical properties of these compounds are poorly known. Although some lattice vibrations have recently been observed by reflection measurements,[1] optical studies have resulted primarily in determinations of the refractive indices and estimates of the intrinsic energy gaps.[2,3] We suggest here a simple method of accurately compensating rather impure *p*-type materials without sample damage. This procedure allows absorption measurements to be made over a wide wave length range. The results are illustrated by measurements on GaSb.

Pulled GaSb crystals normally have a net acceptor density of about $10^{17}/cm^3$. The predominant *p*-type impurity apparently has a distribution coefficient very close to unity. The concentration of this impurity is not appreciably altered by zone refining under a variety of conditions.[4] The spectrum of a typical sample is indicated by curve A in Fig. 1.

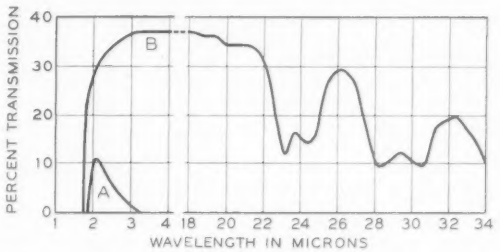


FIG. 1. The room-temperature absorption spectra of GaSb (A) before and (B) after compensation by lithium diffusion. Sample thickness = 0·032 in.

It has been shown that lithium is an interstitial donor[5] which diffuses rapidly[6] in germanium and silicon at relatively low temperatures. The solubility of lithium in the extrinsic range is largely controlled by the hole-electron and ion-pairing equilibria.[7] At temperatures where the intrinsic electron concentration is small, a high degree of compensation of *p*-type samples can be obtained, with any excess lithium precipitating rapidly. Although ion-pair formation has not yet been demonstrated in the III–V compounds, preliminary results indicate a behavior for lithium quite similar to that observed in germanium or silicon. GaSb samples were supersaturated with lithium at 630°C. Any excess lithium was allowed to precipitate at room temperature. The room-temperature resistivity increased from 0·07 to about 1 Ω cm, corresponding to a decrease in acceptor density from $10^{17}/cm^3$ to $5\times10^{15}/cm^3$. Transmission of a typical GaSb sample after diffusion is shown in Fig. 1 by curve B. Lithium diffusion has compensated the sample sufficiently so that the absorption at 34 μ is still less than the minimum absorption of the original material.

Lattice vibrations are observed at 23·3, 24·4, 28·5 and 30·7 μ. Preliminary measurements at

longer wave lengths show additional strong absorptions at approximately 34·5, 40, and 41 μ. It is unlikely that any of the observed bands are ion-pair vibrations which are expected to be much less intense and to occur at longer wave lengths.* These data show two anticipated differences between GaSb and silicon or germanium: (1) the lattice bands of GaSb are more intense than those of silicon or germanium, which to a first approximation are not allowed in the infrared, and (2) the absorption spectrum of GaSb is more complicated because it has lower symmetry than silicon or germanium.

This technique of compensation by lithium diffusion† should permit observation of the lattice vibrations and perhaps of excited states of acceptors in III–V compounds of relatively low purity. Controlled degrees of compensation can also be used to obtain filters for various spectral regions. More detailed investigations of the optical properties of the III–V compounds are in progress.

Acknowledgements—We are indebted to R. H. KAISER and J. A. DITZENBERGER for assistance in sample preparation.

H. J. HROSTOWSKI
C. S. FULLER

Bell Telephone Laboratories, Inc.
Murray Hill, N.J.

REFERENCES

1. PICUS G. S., BURSTEIN E. and HENVIS B. W. *Bull. Amer. Phys. Soc.* **2,** 66 (1957).
2. BRIGGS H. B., CUMMINGS R. F., HROSTOWSKI H. J. and TANENBAUM M. *Phys. Rev.* **93,** 912 (1954).
3. OSWALD F. and SCHADE R. *Z. Naturf.* **9a,** 611 (1954).
4. HROSTOWSKI H. J. Unpublished results.
5. FULLER C. S. and DITZENBERGER J. A. *Phys. Rev.* **91,** 193 (1954).
6. FULLER C. S. and SEVERIENS J. C. *Phys. Rev.* **95,** 21 (1954).
7. REISS H. and FULLER C. S. *J. Metals, N.Y.* **8,** 276 (1956); *J. Phys. Chem. Solids* In press.
8. KOVER F. and QUILLIET A. *C.R. Acad. Sci., Paris* **244,** 1739 (1957).

* Spectra of germanium samples with about 10^{17} ion-pairs/cm³ show little difference from the spectra of pure germanium.

† KOVER and QUILLIET[8] have recently reported compensation of *p*-type AlSb by doping the melt with lithium.

Superconductivity in the Ni–Pd–As system

(*Received* 5 *June* 1957)

RECENTLY, one of us (B. T. M.) and co-workers[1] reported on a number of superconducting compounds, among which were mentioned some solid solutions containing various amounts of nickel, palladium, and arsenic.

$PdAs_2$ has the pyrite structure.[2] It can, as already observed by THOMASSEN,[2] dissolve excess palladium almost to the extent of 50 atomic per cent (PdAs). An X-ray powder photograph of "PdAs" indicates that it has the same lattice constant as $PdAs_2$, 5·982±0·002 Å (agreeing very well with THOMASSEN's value). However, there is present in very small quantity a phase richer in palladium. Even a 60 atomic per cent palladium alloy contains a large proportion of the $PdAs_2$ phase as well as a more palladium-rich phase. A 40 atomic per cent palladium alloy seems to be a pure phase. The lattice constant of the $PdAs_2$ phase remains constant over the whole range.

All of the samples in the Ni–Pd–As system discussed above and in reference (1) were made by solid–solid reaction as follows: Weighed amounts of the reactants were mixed thoroughly, sealed in fused silica tubes, and kept at 1000°C for 24 hr.

The sample represented by $(Ni_{0·75}Pd_{0·25})As$ has mostly the NiAs structure with palladium dissolved in it. There is, however, some of the phase with the pyrite structure ($PdAs_2$, ideal formula) also present. The lattice constants of the NiAs phase increase, whereas, because of the solution of nickel therein, those of the $PdAs_2$ phase decrease. In the samples represented by $(Ni_{0·25}Pd_{0·75})As$, the NiAs phase has the maximum lattice constants $a = 3·66$, $c = 5·03 \pm 0·02$ Å. In the sample represented by $(Ni_{0·12}Pd_{0·88})As$, no NiAs phase is detectable. The X-ray powder photograph indicates the presence of the $PdAs_2$ phase as well as some lines from a more palladium-rich phase. The lattice constant of the $PdAs_2$ phase is a minimum (5·943±0·004 Å), among the compositions studied, for the $(Ni_{0·25}Pd_{0·75})As$, i.e. where the NiAs constants are maximum.

"PdAs" with the pyrite structure is not superconducting above 1·02°K. This is true of the homogeneity range of the $PdAs_2$ structure from 33 to 50 atomic per cent palladium which was studied. Nor

is pure NiAs superconducting above 1·28°K.[3,4] In the measurements of the Ni–Pd–As samples, the highest superconducting transition temperature was 1·6°K, for $(Ni_{0·25}Pd_{0·75})As$,[1] corresponding either to maximum NiAs or minimum $PdAs_2$ lattice constants. The sample which is completely superconducting is that with composition represented by $(Ni_{0·12}Pd_{0·88})As$. This must mean that the $PdAs_2$ phase was made superconducting with addition of nickel. The measurements showed that the remaining samples were only partially superconducting, the amount decreasing with increasing nickel content. Therefore, the situation with regard to superconductivity in the NiAs phase is inconclusive.

In order to confirm the observations with regard to superconductivity in the pyrite phase, we recently prepared new samples of composition similar to the old. However, this time the reaction mixtures were actually melted. The sample represented by $(Ni_{0·12}Pd_{0·88})As$ consisted entirely of single-phase pyrite structure and had a superconducting transition temperature of 1·39°K. The sample represented by $(Ni_{0·25}Pd_{0·75})As$ appeared to be multiphase pyrite, i.e. it had a number of different compositions producing a range of lattice constant.[5] This sample became superconducting at 1·34°K.

It appears that under the proper electronic conditions, the pyrite structure is favorable to superconductivity. There are two other pyrite phases which are superconducting: $RhSe_2$,[5] though the composition which is optimum for superconductivity is $RhSe_{1·75}$,[1] and $RhTe_2$.[1,6]

Acknowledgements—The authors wish to thank E. Corenzwit and V. B. Bala for taking, and Miss B. B. Cetlin for measuring, the photographs used in this investigation.

S. Geller
B. T. Matthias

Bell Telephone Laboratories, Inc.
Murray Hill, N.J.

REFERENCES

1. Matthias B. T., Corenzwit E. and Miller C. E. *Phys. Rev.* **93,** 1415 (1954).
2. Thomassen L. *Z. Phys. Chem.* B **4,** 277 (1929).
3. Matthias B. T. and Hulm J. K. *Phys. Rev.* **87,** 799 (1952).
4. Meissner W., Franz H. and Westerhoff H. *Ann. Phys., Lpz.* **17,** 593 (1933).
5. Geller S. and Cetlin B. B. *Acta Cryst.* **8,** 272 (1955).
6. Geller S. *J. Amer. Chem. Soc.* **77,** 2641 (1955).

CONTENTS LISTS OF VOLUME 2, NOS. 3 AND 4

Volume 2 No. 3

page

T. C. HARMAN, B. PARIS, S. E. MILLER ET H. L. GOERING:
Préparation et quelques propriétés physiques de Bi_2Te_3, Sb_2Te_3 et As_2Te_3 181

MILAN D. FISKE:
Effet d'une contrainte élastique sur la transition de supraconductivité de l'étain et effets de la pression sur les transitions du thallium, du mercure et du tantale 191

A. A. ABRIKOSOV:
Les propriétés magnétiques des alliages supraconducteurs 199

ROBERT E. BEHRINGER:
Raie de résonance nucléaire du cuivre dans les alliages CuMn 209

ERNST SCHLÖMANN:
Théorie de resonances infra-rouges dans les ferrimagnetiques 214

FRANZ R. BROTZEN:
Émission électronique anormale par des surfaces métalliques 221

D. B. FISCHBACH ET A. S. NOWICK:
Sensibilité à la structure de la conductibilité ionique des cristaux NaCl à basse température 226

ARTHUR PASKIN:
Fractions de l'énergie totale et de l'entropie d'ordre magnétique atteintes au point de Curie par les sels antiferromagnétiques 232

K. J. TAUER ET R. J. WEISS:
Une analyse de la chaleur spécifique du manganèse et de sa structure antiferromagnétique. 237

J. BLACK, E. M. CONWELL, L. SEIGLE ET C. W. SPENCER:
Propriétés optiques et électriques de quelques semi-conducteurs de formule $M_2^{V-B}N_3^{VI-B}$. 240

Revue de livre 252

Volume 2 No. 4

WERNER MARTIENSSEN:
Sur les bandes d'excitons des cristaux d'halogénures alcalins 257

S. H. KOENIG ET G. R. GUNTHER-MOHR:
La conductibilité électrique à basse température du germanium de type *n* 268

LEO BREWER ET PETROS ZAVITSANOS:
Étude du systeme Ge–GeO_2 par un appareil à analyse thermique différentielle chauffé par induction 284

LEO BREWER ET FRANK T. GREENE:
Analyse thermique différentielle du system Si–SiO_2 286

M. K. WILKINSON, N. S. GINGRICH ET C. G. SHULL:
La structure magnétique de Mn_2Sb 289

A. KELLER ET A. MARADUDIN:
Diffraction des rayons X par des fibres formées de petits cristaux application de la théorie aux polyamides 301

R. N. SUMMERGRAD ET E. BANKS:
Nouveaux oxydes ferrimagnétiques hexagonaux 312

D. G. THOMAS ET J. J. LANDER:
Conductibilité superficielle produite sur l'oxyde de zinc par le zinc et l'hydrogène . . 318

R. A. ORIANI:
Revue des modèles théoretiques pour les solutions métalliques primaires 327

Lettres aux Editeurs:

G. DE VRIES ET G. W. RATHENAU: La constante de Hall du chrome et la transition à 38°C . 339

MASAHARU NAMBA: Propriétés optiques du carbure de silicium 339

VOL. 4 1958

Volume 2 No. 3

page

Т. Гарман, Б. Парис, С. Е. Миллер и Г. Л. Гёринг :
Приготовление и некоторые физические свойства Bi_2Te_3, Sb_2Te_3, $AsTe_3$. 181

М. Д. Фиске :
Влияние упругой деформации на сверхпроводящий переход в олове и влияние давления на переходы в таллии, ртути и тантале . . . 191

А. А. Абрикосов :
Магнитные свойства сверхпроводящих сплавов 199

Р. Е. Берингер :
Ядерный резонанс меди в сплаве CuMn 209

Е. Шлёман :
Теория инфра-красного резонанса в ферримагнетиках 214

Ф. Р. Бротцен :
Аномальная электронная эмиссия с металлической поверхности . . 221

Д. Б. Фишбах и А. С. Новик :
Структурная чувствительность низко температурной проводимости в кристаллах NaCl 226

А. Паскин :
Доля полной энергии и магнитная энтропия, достигаемые антиферромагнитными солями в точке Кюри 232

К. Дж. Таер и Р. Дж. Вейсс :
Анализ теплоемкости α-марганца и его антиферромагнитная структура . 237

Дж. Блек, Е. М. Конвелл, Л. Зейгле и К. В. Спенсер :
Электрические и оптические свойства некоторых сверхпроводников состава $M_2^{V-B}N_3^{VI-B}$ 240

Рецензия : 252

Volume 2 No. 4

В. Мартинссен :
Об экситонных полосах в кристаллах галогенидов щелочных металлов . 257

С. Г. Кёниг и Г. Р. Гунтер-Мор :
Низкотемпературная электрическая проводимость *n* германия . . 268

Л. Брюер П. Завитсанос :
Исследование системы Ge–GeO_2 в дифференциальном термическом анализаторе с индуктивным нагревом 284

Л. Брюер и Ф. Т. Грин :
Дифференциальный термический анализ системы Si–SiO_2 . . . 286

М. К. Вилкинсон, Н. С. Гингрич и К. Г. Шулл :
Магнитная структура Mn_2Sb 289

А. Келлер и А. Марадудин :
Диффракция X-лучей волокнами, состоящими из маленьких кристаллов : применение теории к полиамидам 301

Р. Н. Зуммерград и Е. Банкс :
Новые гексагональные ферримагнитные окиси 312

Д. Г. Томас и Дж. Ландер :
Поверхностная проводимость, создаваемая в окиси цинка цинком и водородом 318

Р. А. Ориани :
Обзор теоретических моделей первичных металлических растворов . . 327

Письма в редакцию :
Г. де Фриз и Г. В. Ратенау : Постоянная Холла хрома и переход при 38°C . 339
М. Намба : Оптические свойства карбида кремния 339

Volume 2 No. 3

page

T. C. HARMAN, B. PARIS, S. E. MILLER UND H. L. GOERING:
Herstellung und einige physikalische Eigenschaften von Bi_2Te_3, Sb_2Te_3 und As_2Te_3 . 181

MILAN D. FISKE:
Der Einfluss einer elastischen Spannung auf den Übergang zur Supraleitung bei Zinn und die Druckabhängigkeit des Übergangs bei Thallium, Quecksilber und Tantal . . 191

A. A. ABRIKOSOV:
Die magnetischen Eigenschaften von supraleitenden Legierungen 199

ROBERT E. BEHRINGER:
Kernresonanzlinien von Kupfer in CuMn—Legierungen 209

ERNST SCHLÖMANN:
Eine Theorie infraroter Resonanzen in Ferromagnetikas 214

FRANZ R. BROTZEN:
Anomale Elektronenemission von Metalloberflächen 221

D. B. FISCHBACH UND A. S. NOWICK:
Strukturempfindlichkeit der Ionenleitfähigkeit von NaCl—Kristallen bei tiefen Temperaturen 226

ARTHUR PASKIN:
Der Bruchteil der Gesamtenergie- und Entropie wie er bei der Curietemperatur infolge magnetischer Ordnung bei antiferromagnetischen Salzen erhalten wird . . . 232

K. J. TAUER UND R. J. WEISS:
Eine Analyse der spezifischen Wärme des α-Mangans und seine antiferromagnetische Struktur 237

J. BLACK, E. M. CONWELL, L. SEIGLE UND C. W. SPENCER:
Elektrische und optische Eigenschaften einiger $M_2^{V-B}N_3^{VI-B}$-Halbleiter 240

Buchbesprechung 252

Volume 2 No. 4

WERNER MARTIENSSEN:
Über die Excitonenbanden der Alkalihalogenidkristalle 257

S. H. KOENIG UND G. R. GUNTHER-MOHR:
Die elektrische Leitfähigkeit des Germaniums vom *n*-Typ bei tiefen Temperaturen . . 268

LEO BREWER UND PETROS ZAVITSANOS:
Eine Untersuchung des Ge–GeO_2-Systems mit einem induktiv geheizten DTA-Apparat . 284

LEO BREWER UND FRANK T. GREENE:
Eine differentielle thermische Analyse des Si–SiO_2-Systems 286

M. K. WILKINSON, N. S. GINGRICH UND C. G. SHULL:
Die magnetische Struktur des Mn_2Sb 289

A. KELLER UND A. MARADUDIN:
Die Beugung von Röntgenstrahlen durch Fasern die aus kleinen Kristallen bestehen: Anwendung der Theorie auf Polyamide 301

R. N. SUMMERGRAD UND E. BANKS:
Neue hexagonale ferrimagnetische Oxyde 312

D. G. THOMAS UND J. J. LANDER:
Die von Zink und Wasserstoff hervorgerufene Oberflächenleitfähigkeit von Zinkoxyd . . 318

R. A. ORIANI:
Ein Überblick über die theoretischen Modelle für primäre metallische Lösungen . . 327

Briefe an die Herausgeber 339

VOL.
4
1958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 161–176.

ELEKTRONENTHEORETISCHE UNTERSUCHUNGEN ÜBER FEHLSTELLEN IN METALLEN*—V

ALLGEMEINE THEORIE DES EINFLUSSES VON GITTERFEHLSTELLEN AUF DIE TRANSPORTVORGÄNGE IN METALLEN

HELMUT BROSS und ALFRED SEEGER

Max-Planck-Institut für Metallforschung, Stuttgart, und Institut für theoretische und angewandte Physik der Technischen Hochschule Stuttgart

(*Received* 1 *April* 1957)

Zusammenfassung—Die vorliegende Arbeit gibt in möglichst grosser Allgemeinheit die Anwendung des Kohler-Enskogschen Variationsprinzips auf die Streuung von Leitungselektronen durch Gitterfehler in Metallen. Besonderer Wert wird auf die Behandlung der (mit anderen Methoden bis jetzt nicht zugänglichen) Fälle gelegt, in denen der Relaxationszeittensor mehr als eine wesentliche Komponente hat. Das Verfahren gestattet die Berechnung des elektrischen und thermischen Restwiderstandes bei tiefen Temperaturen sowie des elektrischen Widerstandes, des Wärmewiderstandes und der Thermokraft von Fehlstellen oberhalb der Debye-Temperatur. Die Abweichungen von der Matthiessenschen Regel und die Zusammenhänge mit der Anisotropie der Streuer werden ausführlich diskutiert.

Abstract—The present paper treats the application of the Kohler-Enskog variational principle to the scattering of conduction electrons by lattice defects in as general a way as seems possible. Particular emphasis is laid on those cases (which at present cannot be dealt with by other methods) in which the relaxation time tensor has more than one essential component. The technique allows one to calculate the electrical and thermal residual resistivity at low temperatures, and the electrical resistivity, thermal resistivity and thermoelectric power due to defects at temperatures above the Debye-temperature. Deviations from Matthiessen's rule and the connexion with the anisotropy of the scattering centres are discussed in detail.

1. EINLEITUNG

BEIM Studium statischer Gitterfehler in Metallen ist die experimentelle Untersuchung ihres Einflusses auf den elektrischen Widerstand und (in etwas geringerem Masse) auf die Thermokraft und den elektronischen Anteil der Wärmeleitfähigkeit von Wichtigkeit. Die Lösung der damit zusammenhängenden theoretischen Probleme erfordert die Behandlung der gleichzeitigen Streuung der Leitungselektronen durch die thermischen Gitterschwingungen und die statischen Gitterfehler. In ihrer vollen Allgemeinheit sind die dabei auftretenden Fragestellungen so kompliziert, dass bisher nur in einigen Spezialfällen Lösungen der Transportgleichungen gegeben werden konnten, wobei fast immer die Energieflächen im $\mathfrak{k}$-Raum als sphärisch vorausgesetzt werden mussten.

Von MACKENZIE und SONDHEIMER[1] stammt ein auch von anderen Autoren[2,3,4,5,6] verwendetes Verfahren, die zusätzliche Elektronenstreuung durch Gitterfehler bei hohen Temperaturen mit Störungsrechnung zu behandeln. Es wird dabei angenommen, dass (1) die thermische Streuung der Elektronen gross gegen diejenige durch statische Gitterfehler ist und (2) die Relaxationszeit $\tau = \tau(\epsilon)$

* Frühere Mitteilungen: *Z. Phys.* **144**, 637 (1956); **145**, 161 (1956); **146**, 217 (1956); **146**, 242 (1956). Wir benützen die Gelegenheit, um zwei Druckfehler in der dritten der genannten Arbeiten zu verbessern: Auf der rechten Seite vor Gl. (38) ist ein Faktor (-1) einzufügen. Gl. (38b) muss $\xi_2 = \sqrt{[(1-\xi^2)k_F{}^2\rho_0{}^2 - \kappa + \kappa_1 n\,\xi]}$ lauten.

für die thermische Streuung der Elektronen nur von der Elektronenenergie ϵ abhängig ist. Für die thermische Streuung trifft die letztgenannte Annahme oberhalb der Debyetemperatur zu (wo die Elektron-Phonon-Streuung praktisch elastisch, d.h. ohne nennenswerte Energieübertragung erfolgt), sofern Umklappprozesse vernachlässigt werden können.*

SÄENZ[7] hat eine Iterationsmethode für die Behandlung der Transportgleichung angegeben, bei dem zwar noch die eben unter (1) erwähnte Bedingung erfüllt sein muss, die thermische Streuung jedoch nicht durch eine Relaxationszeit $\tau(\epsilon)$ beschreibbar zu sein braucht.

In einer Reihe von Fällen konnten Lösungen gefunden werden, die bei sehr tiefen Temperaturen zutreffen. Es wurde dabei die Annahme gemacht, dass die Streuung durch Gitterschwingungen klein gegen diejenige durch die statischen Gitterfehler und deshalb vernachlässigbar sei. Hierher gehören zunächst isotrop streuende Gitterfehler,[8,9]† bei denen der Tieftemperaturwiderstand mit dem Hochtemperaturwiderstand übereinstimmt. In diesen Fällen ist also die Matthiessensche Regel über die Temperaturunabhängigkeit des Zusatzwiderstandes von statischen Gitterstörungen erfüllt.

SEEGER und STEHLE[6,10] konnten die Tieftemperaturlösungen der Transportgleichungen für zwei Klassen von anisotrop streuenden Gitterfehlern angeben, die die gemeinsame Eigenschaft haben, dass trotz der Anisotropie des Streuvorgangs der Relaxationszeittensor durch *eine* wesentliche (von Null oder Unendlich verschiedene) Komponente charakterisiert werden kann. Es handelt sich dabei einerseits um unendlich lange Streuer mit rotationssymmetrischem Streupotential, wie sie bei manchen Versetzungen in guter Näherung realisiert sind, und andererseits um partiell reflektierende Ebenen, für die ein unbegrenzter Stapelfehler ein Beispiel darstellt. In beiden Fällen wurde gefunden, dass positive Abweichungen von der Matthiessenschen Regel auftreten, also die Zusatzwiderstände (häufig als Restwiderstand bezeichnet) kleiner als diejenigen bei hohen Temperaturen sind.

Die eben erwähnten, positiven Abweichungen von der Matthiessenschen Regel — ebenso wie die nichttrivialen Beiträge von Gitterfehlern zur Widerstandsänderung im Magnetfeld[11,12] — hängen ursächlich mit der Anisotropie der Streuung zusammen. Da man aus diesen Erscheinungen wertvolle Schlüsse über die Natur der Streuer ziehen kann, besteht ein Bedürfnis nach einem systematischen Lösungsverfahren, das insbe-, sondere auch jene Probleme zu behandeln gestattet, in denen man mehr als eine wesentliche Komponente des Relaxationszeittensors hat. Die Anwendungen der hier zu behandelnden allgemeinen Theorie auf zwei solche Probleme, nämlich den elektrischen Widerstand von Doppelleerstellen und denjenigen von aufgespaltenen Versetzungen (= Stapelfehlerbänder, welche von Versetzungen begrenzt sind) wird in späteren Arbeiten gegeben. In der vorliegenden Arbeit wird eine Methode zur Lösung der Boltzmannschen Gleichung beschrieben, mit der man sowohl die Tief- als auch die Hochtemperaturtransportgrössen in beliebiger Genauigkeit bestimmen kann. Ausser der Forderung, dass der Stossmechanismus elastisch sein soll, werden keinerlei weitere Annahmen gemacht. Es können also gleichzeitig sowohl anisotrope als auch isotrope Gitterfehler in beliebigem Verhältnis vorhanden sein.

Der angestrebten Allgemeinheit entsprechend sind die mathematischen Entwicklungen in dieser Arbeit sehr umfangreich. Um dem Leser das Zurechtfinden zu erleichtern, geben wir einen kurzen Überblick über die einzelnen Abschnitte :

In Abschnitt 2 bringen wir eine kurze Zusammenfassung der formalen Transporttheorie und führen die für anisotrope Streuer charakteristische Transporttensoren, die elektrische Leitfähigkeit, den elektronischen Anteil der Wärmeleitfähigkeit und die Thermokraft, ein.

In Abschnitt 3 wird noch ohne spezielle Annahme über den Stossmechanismus und die Form der Energiefläche die statistische Transportgleichung in ein auf KOHLER[13] bezw. ENSKOG[14] zurückgehendes Variationsverfahren umgeschrieben und dieses mit Hilfe des

* Diese Vernachlässigung erscheint oberhalb der Debyetemperatur erlaubt, da man dort die Gitterschwingungen in guter Näherung durch ein Einstein-Modell beschreiben kann, in dem es ja keine Umklappprozesse gibt.

† Wir sprechen von einem isotrop streuenden Gitterfehler, wenn die Übergangswahrscheinlichkeit nur von dem Winkel Θ zwischen den Ausbreitungsvektoren $\mathfrak{k}$ und $\mathfrak{k}'$ des Elektrons vor und nach der Streuung und nicht von der Orientierung dieser Vektoren gegenüber einer ausgezeichneten Richtung des Gitterfehlers abhängt. Ist diese Bedingung nicht erfüllt, so sprechen wir von einem anisotrop streuenden (oder von einem anisotropen) Gitterfehler.

Ritzschen Verfahrens näherungsweise behandelt. Als Vergleichsfunktionen für die Störung der Verteilungsfunktion durch ein äusseres elektrisches Feld und durch Temperaturinhomogenitäten werden im Wellenvektorraum ($\mathfrak{k}$-Raum) Reihenentwicklungen nach Produkten von Kugelflächenfunktionen und Potenzen der Elektronenenergie verwendet. Die Boltzmannsche Stossgleichung geht dadurch in ein System linearer Gleichungen über. Die Transporttensoren sind dann teilweise durch Extremwerte gewisser Integrale bestimmt.

In Abschnitt 4 und allen folgenden untersuchen wir nur noch elastisch gestreute Elektronen. Es wird gezeigt, dass sich bei einem stark entarteten Elektronengas der elektronische Anteil der Wärmeleitfähigkeit und die Thermokraft durch die elektrische Leitfähigkeit bezw. ihre Ableitung nach der Elektronenenergie darstellen lässt.

In Abschnitt 5 wenden wir das für einen beliebigen Streumechanismus entwickelte Variationsverfahren auf *elastische* Streuprozesse an und führen die Matrixelemente $\mathfrak{U}_{\alpha}^{m}(\varepsilon)$ und $T_{\alpha l}^{m o}$ ein.

Ihre Auswertung für ein quasifreies Elektronengas erfolgt im Abschnitt 6.

Im Abschnitt 7 wird das lineare Gleichungssystem unter der Annahme aufgelöst, dass die Winkelabhängigkeit der gestörten Verteilungsfunktion im $\mathfrak{k}$-Raum durch Kugelflächenfunktionen 1. Ordnung genügend genau beschrieben werden kann. In dieser Näherung erhält man für den elektrischen Widerstand und für den elektronischen Beitrag zur Wärmeleitfähigkeit dieselben Ausdrücke, wie sie SÁENZ[7] für hohe Temperaturen abgeleitet hat. Abweichungen von der Matthiessenschen Regel treten also nur dann auf, wenn Kugelfunktionen höherer Ordnung zur Darstellung der gestörten Verteilungsfunktion nötig werden.

In Abschnitt 8 wird der Einfluss der Symmetrie der Gitterfehler auf die Matrixelemente und auf die Auswahl der Vergleichsfunktionen höherer Ordnung untersucht. Wir betrachten dabei folgende Symmetrieverhältnisse :

(a) Isotrop streuende Gitterfehler.
(b) Streuer mit Symmetrieebene und darauf senkrecht stehender zweizähliger Achse.
(c) Rotationssymmetrische Streuer mit und ohne Windungssinn.
(d) Streuer mit kubischer Symmetrie.

In Abschnitt 9 behandeln wir den Fall, dass neben anisotrop streuenden Gitterfehlern auch isotrope Streuer vorhanden sind bezw. Streumechanismen mit isotroper Übergangswahrscheinlichkeit wirken.

In Abschnitt 10 lösen wir das Transportproblem oberhalb der Debyetemperatur, wo sich die Elektron-Phonon Streuung nach dem oben Gesagten durch einen Streumechanismus mit isotroper Übergangswahrscheinlichkeit beschreiben lässt, mit Hilfe des Variationsverfahrens. Gegenüber dem Verfahren von SÁENZ bietet diese Methode den Vorteil, dass der Einfluss der thermischen Streuung auf die Verteilungsfunktion nur bis zu derselben Näherungsstufe betrachtet werden braucht, wie dies für die Behandlung der Zusatzstreuung gewünscht wird, während man beim Iterationsverfahren die exakte Lösung für die thermische Streuung kennen muss. Im Spezialfall eines rotationssymmetrischen Streupotentials werden die Abweichungen von der Matthiessenschen Regel explizit angegeben.

2. FORMALE TRANSPORTTHEORIE

In einem sonst vollkommen idealen Kristallgebiet sei ein Gitterfehler vorhanden. An dem durch diesen Gitterfehler verursachten Störpotential werden die Leitungslektronen gestreut, wodurch selbst am absoluten Nullpunkt ein endlicher elektrischer Widerstand hervorgerufen wird.

Ohne äussere Felder gilt für die Wahrscheinlichkeit, dass ein Elektronenzustand (charakterisiert durch den Ausbreitungsvektor $\mathfrak{k}$) besetzt ist

$$f_0(\mathfrak{k}) = \frac{1}{\exp([\epsilon(\mathfrak{k})-\zeta]/KT)+1}, \tag{1}$$

wobei $\epsilon(\mathfrak{k})$ die Energie der Elektronen mit dem Ausbreitungsvektor $\mathfrak{k}$, ζ die Elektronenenergie an der Fermioberfläche, T die absolute Temperatur und K die Boltzmannsche Konstante bedeutet.

Unter Wirkung äusserer Felder erfährt die Verteilungsfunktion die Änderungen

$$\left(\frac{\partial f}{\partial t}\right)_{\text{Felder}} = \frac{1}{\hbar}\,\mathrm{grad}_{\mathfrak{k}}\epsilon \cdot \mathrm{grad}_{\mathfrak{r}} f - \\ -\frac{e}{\hbar}\cdot(\mathfrak{E}\cdot\mathrm{grad}_{\mathfrak{k}} f \tag{2}$$

(e Betrag der Elementarladung, $\mathfrak{E}$ elektrische Feldstärke).

Im stationären Zustand wird diese Änderung der Verteilungsfunktion infolge der äusseren Felder gerade durch die Streuung der Elektronen an den Gitterfehlern im Gleichgewicht gehalten. Es gilt also

$$\left(\frac{\partial f}{\partial t}\right)_{\text{Felder}} + \left(\frac{\partial f}{\partial t}\right)_{\text{Stösse}} = 0. \tag{3}$$

Ist $W(\mathfrak{k}, \mathfrak{k}')$ die Wahrscheinlichkeit dafür, dass in der Zeiteinheit eines der beiden Elektronen mit dem Vektor $\mathfrak{k}$ durch Wirkung der Gitterfehler in den Zustand $\mathfrak{k}'$ gestreut wird und $W(\mathfrak{k}', \mathfrak{k})$ die Wahrscheinlichkeit für den umgekehrten Vorgang, dann ist

$$\left(\frac{\partial f}{\partial t}\right)_{\text{Stösse}} = \frac{V}{(2\pi)^3}\int \{W(\mathfrak{k}',\mathfrak{k})\cdot f(\mathfrak{k}')[1-f(\mathfrak{k})]- -W(\mathfrak{k},\mathfrak{k}')\cdot f(\mathfrak{k})[1-f(\mathfrak{k}')]\}\, d\mathfrak{k}'. \tag{4}$$

Damit bei Abwesenheit äusserer Felder die zeitliche Änderung der Verteilungsfunktion verschwindet, muss das Prinzip des detaillierten Gleichgewichtes

$$V(\mathfrak{k},\mathfrak{k}') \equiv W(\mathfrak{k}',\mathfrak{k})\cdot[1-f_0(\mathfrak{k})]\cdot f_0(\mathfrak{k}') = W(\mathfrak{k},\mathfrak{k}')[1-f_0(\mathfrak{k}')]\cdot f_0(\mathfrak{k}) \tag{5}$$

erfüllt sein.* Der durch (5) eingeführte Ausdruck $V(\mathfrak{k},\mathfrak{k}')$ ist symmetrisch in $\mathfrak{k}$ und $\mathfrak{k}'$.

Zur Lösung der Gleichungen (2), (3), und (4) macht man allgemein für die Verteilungsfunktion den Ansatz

$$f(\mathfrak{k}) = f_0(\mathfrak{k}) - \Phi(\mathfrak{k})\cdot\frac{\partial f_0}{\partial \epsilon}. \tag{6}$$

In der Näherung einer linearen Transporttheorie[15] erhält man dann folgende Gleichung

$$\frac{1}{\hbar}\frac{\partial f_0}{\partial \epsilon}\cdot \operatorname{grad}_{\mathfrak{k}}\epsilon\{\mathfrak{F}+\epsilon\mathfrak{G}\} = \mathfrak{L}\{\Phi\}, \tag{7}$$

wobei

$$\mathfrak{F} = -e\cdot\mathfrak{E} - T\operatorname{grad}_{\mathrm{r}}\frac{\zeta}{T} \tag{7a}$$

$$\mathfrak{G} = -\frac{1}{T}\operatorname{grad}_{\mathrm{r}} T \tag{7b}$$

und $\mathfrak{L}$ ein linearer Integraloperator ist

$$\mathfrak{L}\{\Psi(\mathfrak{k})\} = \frac{V}{(2\pi)^3}\cdot\frac{\partial f_0(\mathfrak{k})}{\partial\epsilon}\int\frac{f_0(\mathfrak{k}')}{f_0(\mathfrak{k})}W(\mathfrak{k}',\mathfrak{k})\times \times\{\Psi(\mathfrak{k})-\Psi(\mathfrak{k}')\}\, d\mathfrak{k}' \tag{7c}$$

$$= \frac{V}{(2\pi)^3}\int\frac{\partial f_0(\mathfrak{k}')}{\partial\epsilon}\cdot\frac{f_0(\mathfrak{k})}{f_0(\mathfrak{k}')}W(\mathfrak{k},\mathfrak{k}')\times \times\{\Psi(\mathfrak{k})-\Psi(\mathfrak{k}')\}\, d\mathfrak{k}' \tag{7d}$$

Mit dem in (5) eingeführten Ausdruck $V(\mathfrak{k},\mathfrak{k}')$ lässt sich $\mathfrak{L}$ in der folgenden symmetrischen Form schreiben

$$\mathfrak{L}\{\Psi(\mathfrak{k})\} = -\frac{V}{(2\pi)^3}\times \times\frac{1}{KT}\int V(\mathfrak{k},\mathfrak{k}')\{\Psi(\mathfrak{k})-\Psi(\mathfrak{k}')\}\, d\mathfrak{k}'. \tag{7e}$$

Zur Lösung der Integralgleichung (7) machen wir den Ansatz

$$\Phi(\mathfrak{k}) = \mathfrak{F}\cdot\mathfrak{X}^{(1)}(\mathfrak{k}) + \mathfrak{G}\cdot\mathfrak{X}^{(2)}(\mathfrak{k}). \tag{8}$$

Da $\mathfrak{L}$ ein linearer Integraloperator ist und die Vektoren $\mathfrak{F}$ und $\mathfrak{G}$ beliebig sind, zerfällt (7) in die beiden folgenden Integralgleichungen

$$\mathfrak{L}\{\mathfrak{X}^{(a)}(\mathfrak{k})\} = \frac{1}{\hbar a}\cdot\frac{\partial f_0}{\partial\epsilon}\operatorname{grad}_{\mathfrak{k}}\epsilon^a,\ a=1,2. \tag{9}$$

Hat man (9) gelöst, dann ist die elektrische und Wärmestromdichte gegeben durch

$$\mathfrak{j} = -\frac{2}{(2\pi)^3}\cdot\frac{e}{\hbar}\int\operatorname{grad}_{\mathfrak{k}}\epsilon\cdot f(\mathfrak{k})\cdot d\mathfrak{k} = \frac{2}{(2\pi)^3}\frac{e}{\hbar}\int\frac{\partial f_0}{\partial\epsilon}\cdot\operatorname{grad}_{\mathfrak{k}}\epsilon[\mathfrak{F}\cdot\mathfrak{X}^{(1)}(\mathfrak{k})+ +\mathfrak{G}\cdot\mathfrak{X}^{(2)}(\mathfrak{k})]\, d\mathfrak{k} \tag{10a}$$

$$\mathfrak{w} = \frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar}\int\epsilon\operatorname{grad}_{\mathfrak{k}}\epsilon\cdot f(\mathfrak{k})\cdot d\mathfrak{k} = -\frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar}\int\frac{\partial f_0}{\partial\epsilon}\epsilon\cdot\operatorname{grad}_{\mathfrak{k}}\epsilon[\mathfrak{F}\cdot\mathfrak{X}^{(1)}(\mathfrak{k})+ +\mathfrak{G}\cdot\mathfrak{X}^{(2)}(\mathfrak{k})]d\mathfrak{k}. \tag{10b}$$

Zwischen den beiden Vektoren $\mathfrak{F}$ und $\mathfrak{G}$ einerseits und den Vektoren $\mathfrak{j}$ und $\mathfrak{w}$ andererseits besteht ein linearer Zusammenhang. Zur Abkürzung führen wir folgende Tensoren $A^{(\beta)}$ ($\beta = 1, 2, 3$) ein†

$$A^{(1)}_{\iota\kappa} = -\frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar}\int\operatorname{grad}_{\mathfrak{k}_\iota}\epsilon\cdot\frac{\partial f_0}{\partial\epsilon}\mathfrak{X}^{(1)}_\kappa(\mathfrak{k})\, d\mathfrak{k} \tag{11a}$$

$$A^{(2)}_{\iota\kappa} = -\frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar}\int\operatorname{grad}_{\mathfrak{k}_\iota}\epsilon\cdot\frac{\partial f_0}{\partial\epsilon}\mathfrak{X}^{(2)}_\kappa(\mathfrak{k})\, d\mathfrak{k} \tag{11b}$$

* Dies setzt voraus, dass die im Kristall vorkommenden Eigenspannungen keine zu starke Störung der fermischen Verteilungsfunktion zufolge haben.

† Die so definierten Tensoren $A^{(\beta)}$ erfüllen die Onsagerschen Reziprozitätsbeziehungen.

VOL. 4 1958

$$A_{\iota\kappa}^{(3)} = -\frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar}\int \epsilon\,\mathrm{grad}_{\mathfrak{k}_\iota}\epsilon\cdot\frac{\partial f_0}{\partial\epsilon}\mathfrak{X}_\kappa^{(2)}(\mathfrak{k})\,d\mathfrak{k}. \tag{11c}$$

Wie KOHLER[16] gezeigt hat, sind die transponierten Tensoren gegeben durch

$$\tilde{A}^{(1)} = A^{(1)} \tag{12a}$$

$$\tilde{A}^{(3)} = A^{(3)} \tag{12b}$$

$$\tilde{A}_{\iota\kappa}^{(2)} = -\frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar}\int \epsilon\cdot\mathrm{grad}_{\mathfrak{k}_\iota}\epsilon\cdot\frac{\partial f_0}{\partial\epsilon}\cdot\mathfrak{X}_\kappa^{(1)}(\mathfrak{k})\,d\mathfrak{k}. \tag{12c}$$

Zwischen den Tensoren $A^{(\beta)}$ und den üblichen Transportgrössen besteht folgender Zusammenhang:

Elektrische Leitfähigkeit

$$\sigma = e^2\cdot A^{(1)} \tag{13}$$

Wärmeleitfähigkeit

$$\varkappa = T^{-1}(A^{(3)} - \tilde{A}^{(2)}A^{(1)-1}A^{(2)}) \tag{14}$$

Absolute Thermokraft

$$S = -T^{-1}\cdot e^{-1}(A^{(1)-1}A^{(2)} - \zeta\cdot\mathbf{1}). \tag{15}$$

3. DAS VARIATIONSVERFAHREN ZUR LÖSUNG DER TRANSPORTGLEICHUNG BEI EINEM BELIEBIGEN STREUMECHANISMUS

Das eigentliche Problem in der Leitfähigkeitstheorie besteht in der Lösung der Operatorgleichung (9). Eine exakte Lösung dieser Integralgleichung ist nur in wenigen Fällen möglich. Im allgemeinen ist man auf Näherungsverfahren angewiesen, von denen im folgenden das von KOHLER[13] angegebene Variationsverfahren beschrieben werden soll.

Zur Abkürzung werden folgende Bezeichnungen eingeführt: Sind Ψ_1 und Ψ_2 zwei Funktionen von $\mathfrak{k}$, so sei

$$(\Psi_1,\Psi_2) = -\frac{2}{(2\pi)^3}\int \Psi_1\cdot\mathfrak{L}\{\Psi_2\}\,d\mathfrak{k}. \tag{16}$$

Wegen der Symmetrie des Ausdruckes $V(\mathfrak{k},\mathfrak{k}')$ in $\mathfrak{k}$ und $\mathfrak{k}'$ gilt allgemein[16,17]

$$(\Psi_1,\Psi_2) = (\Psi_2,\Psi_1) \tag{17}$$

$$(\Psi,\Psi) \geqslant 0. \tag{18}$$

Das Variationsverfahren besteht nun darin, dass neben den Funktionen $\mathfrak{X}^{(a)}(\mathfrak{k})$, die Lösungen von (9) sind, noch weitere Funktionen $\mathfrak{Y}^{(a)}(\mathfrak{k})$ betrachtet werden, die den folgenden Nebenbedingungen genügen:

1. $(\mathrm{Y}_\iota^{(a)}, \mathrm{Y}_\iota^{(a)})$

$$= -\frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar a}\int\frac{\partial f_0}{\partial\epsilon}\cdot\mathrm{grad}_{\mathfrak{k}_\iota}\epsilon^a\cdot\mathrm{Y}_\iota^{(a)}\,d\mathfrak{k} \equiv (\mathrm{Y}_\iota^{(a)},\mathrm{X}_\iota^{(a)}) \tag{19a}$$

$$\iota = 1,2,3$$

2. $$\int\frac{\partial f_0}{\partial\epsilon}\mathrm{Y}_\iota^{(a)}\cdot d\mathfrak{k} = 0 \tag{19b}$$

3. $$\int\frac{\partial f_0}{\partial\epsilon}\epsilon\cdot\mathrm{Y}_\iota^{(a)}\cdot d\mathfrak{k} = 0. \tag{19c}$$

Wegen (18) gilt nun

$$(\mathrm{X}_\iota^{(a)} - \mathrm{Y}_\iota^{(a)}, \mathrm{X}_\iota^{(a)} - \mathrm{Y}_\iota^{(a)}) \geqslant 0.$$

Berücksichtigen wir noch (17) und die Nebenbedingung (19a), dann ergibt sich

$$(\mathrm{X}_\iota^{(a)},\mathrm{X}_\iota^{(a)}) \geqslant (\mathrm{Y}_\iota^{(a)},\mathrm{Y}_\iota^{(a)}) \tag{20}$$

$$\iota = 1,2,3$$

$$a = 1,2.$$

Unter allen Funktionen $\mathrm{Y}_\iota^{(a)}(\mathfrak{k})$, die den Nebenbedingungen (19a,b,c) genügen, machen die Lösungen $\mathrm{X}_\iota^{(a)}(\mathfrak{k})$ der Integralgleichung (9) den Ausdruck $(\mathrm{Y}_\iota^{(a)}, \mathrm{Y}_\iota^{(a)})$ zu einem Maximum.

Zur Lösung des Extremalproblems (20) mit den Nebenbedingungen (19a,b,c) benützen wir das Ritzsche Verfahren.

Im allgemeinen ist die Verteilungsfunktion $f(\mathfrak{k})$ und damit auch $\mathfrak{X}^{(a)}(\mathfrak{k})$ von drei Variablen abhängig. Zur Festlegung von $\mathfrak{k}$ benützt man am besten ein sphärisches Polarkoordinatensystem (ϑ,ϕ) im $\mathfrak{k}$-Raum und als dritte Variable die Elektronenenergie ϵ:

$$\begin{aligned} k_x &= k(\epsilon,\vartheta,\phi)\cdot\sin\vartheta\cdot\cos\phi \\ k_y &= k(\epsilon,\vartheta,\phi)\cdot\sin\vartheta\cdot\sin\phi \\ k_z &= k(\epsilon,\vartheta,\phi)\cdot\cos\vartheta. \end{aligned} \tag{21}$$

Für die Vergleichsfunktionen machen wir den Ansatz*

$$\mathrm{Y}_{\iota}^{(a;p,q)} = \sum_{r=0}^{p}\sum_{n=0}^{q}\sum_{m=-n}^{n} \mathrm{Y}_{\iota,\,r\;\;n}^{(a;p,q)m} \times$$

$$\times\left(\frac{\epsilon-\zeta}{\zeta}\right)^{r}\cdot \Pi_n^m(\cos\vartheta)\cdot e^{im\phi}. \tag{22}$$

Wir suchen also die richtige Verteilungsfunktion $\mathrm{X}_{\iota}^{(a)}(\mathfrak{k})$ durch die zweidimensionale Funktionenfolge $\mathrm{Y}_{\iota}^{(a,p,q)}$ mit $p=0,1,2,\ldots$ und $q=0,1,2,\ldots$ anzunähern. Dabei entspricht wachsendes p einer immer besser werdenden Näherung für die Energieabhängigkeit, während wachsendes q eine bessere Näherung für die Winkelabhängigkeit bedeutet.

Die Nebenbedingungen (19b) und (19c) lassen sich erfüllen durch $\mathrm{Y}_{\iota\;\;r\;\;0}^{(a;p,q)\,0}=0$; die Summation über n beginnt also erst mit $n=1$. Die Entwicklungskoeffizienten kann man nun mit den üblichen Verfahren zur Lösung gewöhnlicher Extremalprobleme bestimmen. Wir erhalten

$$(\mathrm{Y}_{\iota}^{(a;p,q)},\,\mathrm{Y}_{\iota}^{(a;p,q)})$$

$$=\sum_{r=0}^{p}\sum_{s=0}^{p}\sum_{n=1}^{q}\sum_{l=1}^{q}\sum_{m=-n}^{n}\sum_{o=-l}^{l} \mathrm{Y}_{\iota,\;\;r\;\;n}^{(a;p,q)m}\times$$

$$\times\,\mathrm{Y}_{\iota\;\;s\;\;l}^{(a;p,q)\,o}\cdot {}_{rs}M_{n\;\;l}^{m\;\;o}, \tag{23a}$$

wobei

$${}_{rs}M_{n\;\;l}^{m\;\;o}=\left[\left(\frac{\epsilon-\zeta}{\zeta}\right)^{r}\cdot\Pi_n^m\cdot e^{im\phi};\left(\frac{\epsilon-\zeta}{\zeta}\right)^{s}\Pi_l^o\cdot e^{io\phi}\right] \tag{23b}$$

* $\Pi_n^m(\cos\vartheta)$ sind normierte Kugelfunktionen

$$\Pi_n^m(\cos\vartheta)$$

$$=\left[\frac{2n+1}{2}\cdot\frac{(n-m)!}{(n+m)!}\right]^{1/2}\cdot P_n^m(\cos\vartheta)$$

$$=(-1)^m\cdot\Pi_n^{-m}(\cos\vartheta)$$

$$\int_0^{\pi}\Pi_n^m(\cos\vartheta)\cdot\Pi_n^m(\cos\vartheta)\sin\vartheta\,d\vartheta=1.$$

ist. Die Nebenbedingung (19a) lautet

$$(\mathrm{Y}_{\iota}^{(a;p,q)},\,\mathrm{Y}_{\iota}^{(a;p,q)}) \tag{24a}$$

$$=\sum_{r=0}^{p}\sum_{n=1}^{q}\sum_{m=-n}^{n} \mathrm{Y}_{\iota,\;\;r\;\;n}^{(a;p,q)m}\cdot {}_{r}N_{\iota\;\;n}^{(a)\,m}$$

mit

$${}_{r}\mathfrak{N}_{n}^{(a)m}=-\frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar}\cdot\frac{1}{a}\int\frac{\partial f_0}{\partial\epsilon}\cdot\mathrm{grad}_{\mathfrak{k}}\epsilon^{a}\left(\frac{\epsilon-\zeta}{\zeta}\right)^{r}\times$$

$$\times\,\Pi_n^m(\cos\vartheta)\cdot e^{im\phi}\cdot d\mathfrak{k}. \tag{24b}$$

Die Extremalforderungen (20) lassen sich mit Hilfe eines Lagrangeschen Faktors wie folgt schreiben:

$$2\sum_{s=0}^{v}\sum_{l=1}^{q}\sum_{o=-l}^{l} \mathrm{Y}_{\iota,\;\;s\;\;l}^{(a;p,q)\,o}\cdot {}_{rs}M_{n\;\;l}^{m\;\;o}+\lambda\cdot {}_{r}N_{\iota\;\;n}^{(a)\,m}=0$$

$$\begin{aligned} &n=1,2,\ldots,q\\ &-n<m<n\\ &r=0,1,2,\ldots,p. \end{aligned} \tag{25}$$

Zur Bestimmung des Multiplikators λ multiplizieren wir (25) mit $\mathrm{Y}_{\iota,\;\;r\;\;n}^{(a,p;q)m}$ und summieren über r, n und m. Aus (23a) und (24a) folgt $\lambda=-2$. Die Entwicklungskoeffizienten $\mathrm{Y}_{\iota,\;\;r\;\;n}^{(a,p;q)m}$ sind also durch folgendes lineares Gleichungssystem festgelegt:

$$\sum_{s=0}^{p}\sum_{l=1}^{q}\sum_{o=-l}^{l} \mathrm{Y}_{\iota,\;\;s\;\;l}^{(a;p,q)\,o}\cdot {}_{rs}M_{n\;\;l}^{m\;\;o}={}_{r}N_{\iota\;\;n}^{(a)\,m}$$

$$\begin{aligned} &n=1,2,\ldots,q\\ &-n<m<n\\ &r=0,1,2,\ldots,p\\ &a=1,2. \end{aligned} \tag{26}$$

Da die ${}_{rs}M_{nl}^{mo}$ in Gleichung (23a) Koeffizienten einer positiv definiten quadratischen Form sind, ist die Determinante $|M_{nl}^{mo}|$ positiv. Man kann also das lineare Gleichungssystem (26) immer auflösen.

Anstelle der eingeführten Tensoren $A^{(\beta)}$ ($\beta=1,2,3$) betrachten wir die Näherungstensoren $A^{(\beta;p,q)}$, die den Vergleichsfunktionen $\mathfrak{Y}^{(\beta;p,q)}$ entsprechen und die durch

$$A_{\iota\kappa}^{(1;p,q)}=-\frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar}\int\mathrm{grad}_{\mathfrak{k}_\iota}\epsilon\cdot\frac{\partial f_0}{\partial\epsilon}\cdot\mathrm{Y}_{\kappa}^{(1;p,q)}\cdot d\mathfrak{k}$$

$$\equiv(\mathrm{X}_{\iota}^{(1)}(\mathfrak{k}),\,\mathrm{Y}_{\kappa}^{(1;p,q)}(\mathfrak{k})) \tag{27a}$$

VOL. 4 1958

$$A_{\iota\kappa}^{(2;p,q)} = -\frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar}\int \operatorname{grad}_{\mathfrak{k}_\iota}\epsilon\cdot\frac{\partial f_0}{\partial\epsilon}\cdot Y_\kappa^{(2;p,q)}\cdot d\mathfrak{k} \equiv (X_\iota^{(1)}(\mathfrak{k}), Y_\kappa^{(2;p,q)}(\mathfrak{k})) \tag{27b}$$

$$A_{\iota\kappa}^{(3;p,q)} = -\frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar}\int \epsilon\operatorname{grad}_{\mathfrak{k}_\iota}\epsilon\cdot\frac{\partial f_0}{\partial\epsilon}\cdot Y_\kappa^{(2;p,q)}\cdot d\mathfrak{k} \equiv (X_\iota^{(2)}(\mathfrak{k}), Y_\kappa^{(2;p,q)}(\mathfrak{k})) \tag{27c}$$

definiert sind.*

Damit sind alle Näherungswerte (p, q)-ter Näherung für die Transportgrössen durch Integrale über die Vergleichsfunktionen (p, q)-ter Näherung dargestellt. Wie bei KOHLER[13] lassen sich die Ausdrücke für die Transportgrössen (p, q)-ter Näherung in äusserst eleganter Weise durch Determinanten darstellen. Insbesondere kann man durch mehrmaliges Anwenden des Sylvesterschen Satzes[18] über Superdeterminanten zeigen, dass die elektrische und Wärmeleitfähigkeit mit wachsendem p und q monoton zunehmen. Auf eine Ableitung für den allgemeinen Fall soll verzichtet werden, da die auftretenden Determinanten sehr umfangreich werden.

4. DIE TRANSPORTTHEORIE BEI ELASTISCHER STREUUNG DER ELEKTRONEN

Elektrische Leitfähigkeit, Wärmeleitfähigkeit und Thermokraft bei einem stark entarteten Elektronengas

Bei statischen Gitterfehlern, die zeitunabhängige Störpotentiale hervorrufen, erfolgt die Streuung der Elektronen elastisch:

$$W(\mathfrak{k}, \mathfrak{k}') = w(\mathfrak{k}, \mathfrak{k}')\cdot\delta(\epsilon(\mathfrak{k})-\epsilon(\mathfrak{k}')). \tag{28}$$

Aus dem Prinzip des detaillierten Gleichgewichtes folgt, dass die Verteilungsfunktion $f(\mathfrak{k})$ bei Abwesenheit äusserer Felder nur dann in die Fermische Verteilungsfunktion f_0 übergeht, wenn die Übergangswahrscheinlichkeit symmetrisch in $\mathfrak{k}$ und $\mathfrak{k}'$ ist.

$$W(\mathfrak{k}, \mathfrak{k}') = W(\mathfrak{k}', \mathfrak{k}). \tag{29}$$

* Wegen der Nebenbedingung (19a) sind die Elemente auf der Hauptdiagonale von $A^{(1;p,q)}$ bezw. $A^{(3;p,q)}$ durch die Extremwerte von $(Y_\iota^{(1;p,q)}\ Y_\iota^{(1;p,q)})$ bezw. $(Y_\iota^{(2;p,q)}), Y_\iota^{(2;p,q)})$ bestimmt.

Diese Bedingung soll im folgenden immer erfüllt sein. Die Definitionsgleichung für den linearen Integraloperator $\mathfrak{L}$ lautet dann

$$\mathfrak{L}\{\Psi(\mathfrak{k})\} = \frac{V}{(2\pi)^3}\cdot\frac{\partial f_0}{\partial\epsilon}\int W(\mathfrak{k}, \mathfrak{k}')\{\Psi(\mathfrak{k})-\Psi(\mathfrak{k}')\}\, d\mathfrak{k}'. \tag{30}$$

Da bei elastischer Streuung der Elektronen die Energie erhalten bleibt, ist jede Funktion von ϵ mit dem Operator $\mathfrak{L}$ vertauschbar. Aus (9) folgt

$$\mathfrak{X}^{(2)}(\mathfrak{k}) = \epsilon\cdot\mathfrak{X}^{(1)}(\mathfrak{k}). \tag{31}$$

Für den Tensor $\mathbf{A}^{(\beta)}$ gilt allgemein

$$A_{\iota\kappa}^{(\beta)} = -\frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar}\int \epsilon^{\beta-1}\cdot\operatorname{grad}_{\mathfrak{k}_\iota}\epsilon\cdot\frac{\partial f_0}{\partial\epsilon}X_\kappa^{(1)}\cdot d\mathfrak{k}. \tag{32}$$

Wir brauchen also bloss die Funktion $\mathfrak{X}^{(1)}(\mathfrak{k})$ zu bestimmen. Wenn keine Verwechslung möglich ist, lassen wir bei $\mathfrak{X}^{(1)}(\mathfrak{k})$ in Zukunft den oberen Index (1) weg.

Bei einem stark entarteten Elektronengas bekommt man besonders einfache Verhältnisse: Wärmeleitfähigkeit und Thermokraft lassen sich nämlich durch die elektrische Leitfähigkeit bezw. ihre Ableitung nach der Elektronenenergie darstellen. Um dies zu zeigen, zerlegen wir die Integration über den $\mathfrak{k}$-Raum in ein Oberflächenintegral und ein Integral über die Elektronenenergie ε

$$A_{\iota\kappa}^{(\beta)} = -\frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar}\int \epsilon^{\beta-1}\frac{\partial f_0}{\partial\epsilon}\cdot\int\frac{\partial t}{\partial\epsilon}k^2\cdot\operatorname{grad}_{\mathfrak{k}_\iota}\epsilon\times \times X_\kappa(\epsilon, \vartheta, \phi)\, d\omega_{\mathfrak{I}}\cdot d\epsilon. \tag{33}$$

Führen wir die Integration über ε bis zu quadratischen Gliedern im Entartungsparameter KT/ζ mit Hilfe der Sommerfeldschen Beziehung† durch, so erhalten wir

$$A^{(\beta)} = \left[\epsilon^{\beta-1}\cdot\boldsymbol{\chi}(\epsilon)+\frac{\pi}{6}(KT)^2\frac{d^2}{d\epsilon^2}(\epsilon^{\beta-1}\boldsymbol{\chi}(\epsilon))\right]_{\epsilon=\zeta}^{*}, \tag{34}$$

† Bei einer beliebigen Funktion $g(\varepsilon)$ gilt allgemein

$$\int_0^\infty g(\epsilon)\frac{\partial f_0}{\partial\epsilon}\,d\epsilon = -\left[g(\epsilon)+\frac{\pi^2}{6}(KT)^2\frac{d^2g}{d\epsilon^2}\right]_{\epsilon=\zeta}$$

SOMMERFELD A. und BETHE H. *Handbuch der Physik* XXIV, 2 Berlin (1933).

wobei der Tensor $\boldsymbol{\chi}(\epsilon)$ durch folgendes Oberflächenintegral dargestellt ist :

$$\chi_{\iota\kappa}(\epsilon) = \frac{2}{(2\pi)^3}\cdot\frac{1}{\hbar}\int \frac{\partial k}{\partial \epsilon}k^2\cdot \mathrm{grad}_{\mathfrak{k}\iota}\epsilon\cdot X_\kappa(\epsilon,\vartheta,\phi)\,d\omega_{\mathfrak{k}}. \tag{35}$$

Für die üblichen Transportgrössen ergibt sich nach kurzer Zwischenrechnung,* wenn wir bis zu den ersten im Entartungsparameter nicht verschwindenden Gliedern gehen

$$\boldsymbol{\sigma} = \epsilon^2\cdot\boldsymbol{\chi}(\zeta) \tag{36}$$

$$\boldsymbol{\varkappa} = \frac{\pi^2}{3}K^2\cdot T\boldsymbol{\chi}(\xi) \equiv \frac{\pi^2}{3}\cdot\frac{K^2}{e^2}T\boldsymbol{\sigma} \tag{37}$$

$$\boldsymbol{S} = -\frac{\pi^2}{3}\cdot\frac{K^2}{e}T\left[\boldsymbol{\chi}^{-1}(\epsilon)\frac{d}{d\epsilon}\boldsymbol{\chi}(\epsilon)\right]_{\epsilon=\zeta} \tag{38a}$$

$$\boldsymbol{S} = -\frac{\pi^2}{3}\cdot\frac{K^2}{e}T\left[\boldsymbol{\sigma}^{-1}(\epsilon)\frac{d}{d\epsilon}\boldsymbol{\sigma}(\epsilon)\right]_{\epsilon=\zeta}. \tag{38b}$$

Bei einem stark entarteten Elektronengas ist demnach für einen beliebigen elastischen Stossmechanismus — also bei sehr tiefen Temperaturen und oberhalb der Debyetemperatur — und für jede beliebige Form der Energiefläche das Wiedemann-Franzsche Gesetz gültig, was schon von KOHLER[16] in derselben Allgemeinheit gezeigt wurde.

Aus den Gleichungen (37) und (38b) ist ferner noch ersichtlich, dass sich unter den gemachten Annahmen alle Transportgrössen als Funktionen der elektrischen Leitfähigkeit darstellen lassen.

Die Beziehungen (37) und (38) sind schon bei MOTT und JONES[19] unter der Annahme einer allgemeinen Relaxationszeit $\tau(\mathfrak{k})$ für skalare Leitfähigkeitsgrössen abgeleitet worden. Sie hätten sich auch auf den Fall tensorieller Grössen übertragen lassen, wenn man statt der isotropen Relaxationszeit nach MACKENZIE und SONDHEIMER[1] sowie SEEGER und STEHLE[6] den Relaxationszeittensor $\boldsymbol{\tau}(\mathfrak{k})$ eingeführt hätte. Für auf Hauptachsen transformierte Tensoren findet sich Beziehung (38b) ohne Begründung bei HIRONE und ADACHI.[4]

5. DAS VARIATIONSPROBLEM BEI ELASTISCHER STREUUNG DER ELEKTRONEN

Da bei elastischer Streuung der Elektronen jede beliebige Funktion der Energie ϵ mit dem Integraloperator $\mathfrak{L}$ vertauschbar ist, stellt die Operatorgleichung (9) nur für die beiden Variablen ϑ und ϕ eine Integralgleichung dar, während sie für die Variable ϵ bloss eine algebraische Gleichung ist. Wir könnten nun, in ähnlicher Weise wie wir es in Abschnitt 3 für den gesamten $\mathfrak{k}$-Raum durchführten, ein Extremalprinzip für eine Fläche $\epsilon = \text{const.}$ entwickeln. Diesen Weg wollen wir nicht gehen, sondern wir versuchen, das elastische Streuproblem mit dem allgemeinen Verfahren zu lösen. Dazu führen wir folgende Oberflächenintegrale ein

$$T^{mo}_{nl}(\epsilon) = \frac{(-1)^m}{2\pi}\int \frac{\partial k}{\partial\epsilon}k^2\cdot\Pi^m_n(\cos\vartheta)\times \\ \times e^{im\phi}\left(\frac{\partial f_0}{\partial\epsilon}\right)^{-1}\mathfrak{L}\{\Pi^o e^{io\phi}\}\,d\omega_{\mathfrak{k}} \tag{39}$$

$$\mathfrak{U}^m_n(\epsilon) = \frac{(-1)^m}{2\pi}\cdot\frac{1}{\hbar}\int\frac{\partial k}{\partial\epsilon}k^2\times \\ \times\mathrm{grad}_{\mathfrak{k}}\epsilon\cdot\Pi^m_n(\cos\vartheta)\cdot e^{im\phi}\,d\omega_{\mathfrak{k}}. \tag{40}$$

Das lineare Gleichungssystem (26) geht dann über in

$$\frac{2}{(2\pi)^2}(-1)^m\int_0^\infty\left[\sum_{s=0}^{p}\sum_{l=1}^{q}\sum_{o=-l}^{l}Y^{(p,q)\,o}_{\iota,s\quad l}\times\right. \\ \left.\times T^{m\,o}_{n\,l}(\epsilon)\left(\frac{\epsilon-\zeta}{\zeta}\right)^{r+s}\frac{\partial f_0}{\partial\epsilon}\right]d\epsilon \tag{41}$$

$$= \frac{2}{(2\pi)^2}(-1)^m\int_0^\infty\frac{\partial f_0}{\partial\epsilon}\left(\frac{\epsilon-\zeta}{\zeta}\right)^r U^m_{\iota n}(\epsilon)\,d\epsilon.$$

Statt das Extremalprinzip für jede Fläche $\epsilon = \text{const.}$ aufzustellen, kann man fordern, dass die Integranden auf beiden Seiten von Gleichung (41) dieselben Funktionen von ϵ sind. Dies führt auf

$$\sum_{l=1}^{q}\sum_{o=-l}^{l}Y^{(q)\,o}_{\iota\quad l}(\epsilon)\cdot T^{m\,o}_{n\,l}(\epsilon) \\ = U^m_{\iota n}(\epsilon)\quad \begin{matrix} n=1,2,\ldots,\mathfrak{q} \\ -n<m<n, \end{matrix} \tag{42}$$

* Ist der Tensor $\mathbf{a}$ gross gegenüber dem Tensor $\mathbf{b}$, dann lässt sich der reziproke Tensor $(\mathbf{a}+\mathbf{b})^{-1}$ näherungsweise darstellen durch

$$(\boldsymbol{a}+\boldsymbol{b})^{-1} = \boldsymbol{a}^{-1}-\boldsymbol{a}^{-1}\boldsymbol{b}\boldsymbol{a}^{-1}+\boldsymbol{a}^{-1}\boldsymbol{b}\boldsymbol{a}^{-1}\boldsymbol{b}\boldsymbol{a}^{-1}-\ldots+$$

VOL. 4 1958

wobei zur Abkürzung eingeführt ist

$$Y^{(q)\,o}_{\iota}(\epsilon) = \sum_{s=0}^{p} Y^{(p,q)\,o}_{s\ \iota}(\epsilon)\left(\frac{\epsilon-\zeta}{\zeta}\right)^{s}. \tag{43}$$

Da der Parameter p nur in Gleichung (43) vorkommt, kann man $p \to \infty$ setzen, wodurch die durch Gleichung (43) definierten Entwicklungskoeffizienten unabhängig vom Näherungsgrad in der Energieabhängigkeit werden.

6. BERECHNUNG DES MATRIXELEMENTS $\mathfrak{U}^m_n(\varepsilon)$ BEI EINEM QUASIFREIEN ELEKTRONENGAS

Wir haben das bisher beschriebene Verfahren zur Lösung der Boltzmannschen Gleichung bei Anwesenheit von Gitterfehlern möglichst allgemein gehalten und daher keine Annahmen über die Abhängigkeit der Elektronenenergie $\epsilon(\mathfrak{k})$ vom Ausbreitungsvektor $\mathfrak{k}$ gemacht. Von nun an werden die Elektronen als quasifrei mit der skalaren effektiven Masse m^* behandelt.* Zwischen der Ein elektronenenergie $\epsilon(\mathfrak{k})$ und dem Wellenvektor $\mathfrak{k}$ besteht dann folgender Zusammenhang†

$$\epsilon(\mathfrak{k}) = \frac{\hbar^2}{2m^*}\mathfrak{k}^2 \equiv \frac{\hbar^2}{2m^*}k^2, \tag{44}$$

so dass

$$\operatorname{grad}_{\mathfrak{k}}\epsilon = \frac{d\epsilon}{dk}\operatorname{grad}_{\mathfrak{k}}k = \frac{d\epsilon}{dk}\cdot\frac{\mathfrak{k}}{k} \tag{45}$$

gilt.

Berücksichtigen wir noch, dass sich die kartesischen Komponenten des Einheitsvektors $\mathfrak{k}^0$ durch Kugelfunktionen in der Form

$$k^0_x = \frac{1}{\sqrt{3}}[\Pi^1_1(\cos\vartheta)\cdot e^{i\phi} - \Pi^{-1}_1(\cos\vartheta)\cdot e^{-i\phi}] \tag{46a}$$

$$k^0_y = \frac{1}{\sqrt{3}}\cdot\frac{1}{i}[\Pi^1_1(\cos\vartheta)e^{i\phi} + \Pi^{-1}_1(\cos\vartheta)\cdot e^{-i\phi}] \tag{46b}$$

$$k^0_z = \sqrt{\left(\frac{2}{3}\right)}\,\Pi^0_1(\cos\vartheta) \tag{46c}$$

darstellen lassen, so erkennen wir, dass $\mathfrak{U}^m_n(\epsilon)$ nur dann von Null verschieden ist, wenn $n = 1$ und $m = 0$ oder 1 ist. Für die von Null verschiedenen Matrixelemente ergibt sich aus (40) und (45)

$$\mathfrak{U}^{-1}_1(\epsilon) = \frac{1}{\sqrt{3}}\cdot\frac{k^2}{\hbar}(\mathfrak{e}_1 - i\mathfrak{e}_2), \tag{47a}$$

$$\mathfrak{U}^{1}_1(\epsilon) = \frac{1}{\sqrt{3}}\cdot\frac{k^2}{\hbar}(-\mathfrak{e}_1 - i\mathfrak{e}_2), \tag{47b}$$

$$\mathfrak{U}^{0}_1(\epsilon) = \sqrt{\left(\frac{2}{3}\right)}\cdot\frac{k^2}{\hbar}\mathfrak{e}_3, \tag{47c}$$

wobei $\mathfrak{e}_1$, $\mathfrak{e}_2$ und $\mathfrak{e}_3$ die Einheitsvektoren eines kartesischen Koordinatensystems sind.

7. DIE NÄHERUNG ERSTER ORDNUNG IN DER WINKELABHÄNGIGKEIT

In den Beziehungen für die Leitfähigkeitstensoren kommen bei einem quasifreien Elektronengas nur die Entwicklungskoeffizienten der Kugelflächenfunktionen erster Ordnung vor, während die übrigen Entwicklungskoeffizienten die Leitfähigkeitstensoren nur mittelbar über die Gleichung (42) beeinflussen können. Da ausserdem die exakte Lösung der Transportgleichung bei isotrop streuenden Gitterfehlern bloss Kugelfunktionen erster Ordnung enthält, wird auch bei anisotropen Übergangswahrscheinlichkeiten der Hauptbeitrag zu den Leitfähigkeitstensoren von der ersten Näherung in der Winkelabhängigkeit herrühren. Man erhält mit dieser Näherung jedoch nicht die typisch anisotropen Effekte wie Abweichungen von der Matthiessenschen Regel und Widerstandsänderung im Magnetfeld.

Die Näherung erster Ordnung in der Winkelabhängigkeit lässt sich besonders einfach schreiben, wenn man in (22) zur reellen Schreibweise übergeht und

$$Y^{(1)}_{\iota}(\epsilon) = \sum_{\nu=1}^{3} y^{(1)}_{\iota\nu}(\epsilon)k^0_{\nu} \quad \iota = 1,2,3 \tag{48}$$

* Die weiteren Ergebnisse dieser Arbeit sind damit im wesentlichen auf die einwertigen Metalle und ferromagnetisches Nickel anwendbar.

† In den folgenden Abschnitten benützen wir bloss die Tatsache, dass $\varepsilon(\mathfrak{k})$ eine Funktion vom Betrag des Ausbreitungsvektors $|\mathfrak{k}|$ ist.

setzt.* Auf diese Vergleichsfunktionen kann man nun das in Abschnitt 3 bezw. 5 beschriebene Variationsverfahren anwenden. Für die Entwicklungskoeffizienten $y_{\iota\mu}^{(1)}(\epsilon)$ erhält man folgendes lineare Gleichungssystem

$$\sum_{\mu=1}^{3} y_{\iota\mu}^{(1)}(\epsilon)\, t_{\nu\mu}(\epsilon) = u_{\iota\nu}(\epsilon) \qquad \begin{matrix}\nu = 1,2,3\\ \iota = 1,2,3,\end{matrix} \tag{49}$$

wobei wir zur Abkürzung die Matrixelemente

$$t_{\nu\mu}(\epsilon) = \frac{1}{2\pi}\int k_\nu^0 \left(\frac{\partial f_0}{\partial \epsilon}\right)^{-1} \mathfrak{Q}\{k_\mu^0\} k^2 \frac{dk}{d\epsilon}\, d\omega_{\mathfrak{k}} \tag{50}$$

$$u_{\iota\nu}(\epsilon) = \frac{1}{2\pi}\cdot\frac{1}{\hbar}\int \mathrm{grad}_{\mathfrak{k}_\iota}\epsilon \cdot k_\nu^0 \cdot k^2 \frac{dk}{d\epsilon}\, d\omega_{\mathfrak{k}} \tag{51}$$

eingeführt haben. Wegen (45) ergibt sich für

$$u_{\iota\nu}(\epsilon) = \frac{2}{3}\cdot\frac{k^2}{\hbar}\,\delta_{\iota,\nu}. \tag{52}$$

Um die Tieftemperaturtransportgrössen mit den Hochtemperaturgrössen vergleichen zu können, führen wir den von SÁENZ[7] (Gleichung 3, 22c) definierten Tensor

$$\tau_{\nu\mu}^{-1} = \frac{3}{4\pi}\int k_\nu^0 \left(\frac{\partial f_0}{\partial \epsilon}\right)^{-1} \mathfrak{Q}\{k_\mu^0\}\, d\omega_{\mathfrak{k}} \tag{53}$$

ein.† Dann ist

$$t_{\nu\mu} = \frac{2}{3}k^2\frac{dk}{d\epsilon}\tau_{\nu\mu}^{-1}. \tag{54}$$

Wegen (29) ist $t_{\nu\mu}$ eine in ν und μ symmetrische Matrix. Ohne Einschränkung der Allgemeinheit können wir erreichen, dass durch eine lineare Transformation des $\mathfrak{k}$-Raums $t_{\nu\mu}$ auf Hauptachsen transformiert ist. Gleichung (49) nimmt dann die einfache Form

$$y_{\iota\nu}^{(1)}(\epsilon) t_{\nu\nu} \quad (\epsilon) = u_{\iota\nu} \tag{55}$$

an. Berücksichtigen wir noch (52), so können wir dieses Gleichungssystem erfüllen, wenn wir für $\iota \neq \nu$ den Entwicklungskoeffizienten $y_{\iota\nu}^{(1)}$ gleich Null setzen. Für $\iota = \nu$ erhalten wir nach Auflösung des linearen Gleichungssystems

$$y_{\iota\iota}^{(1)}(\epsilon) = \frac{1}{\hbar}\cdot\frac{d\epsilon}{dk}\tau_{\iota\iota}(\epsilon). \tag{56}$$

Für die Tensoren $\mathbf{A}^{(\beta)}$ $(\beta = 1, 2, 3)$ ergibt sich dann nach kurzer Zwischenrechnung, wenn man nur bis zu quadratischen Gliedern im Entartungsparameter geht,

$$A^{(\beta;1)} = \frac{1}{3\pi^2}\cdot\frac{1}{\hbar^2}\cdot\frac{1}{\beta}\left[k^2\frac{d\epsilon^\beta}{dk}\boldsymbol{\tau} + \frac{\pi^2}{6}(KT)^2\frac{d^2}{d\epsilon^2}\left(k^2\frac{d\epsilon^\beta}{dk}\boldsymbol{\tau}\right)\right]_{\epsilon=\zeta}. \tag{57}$$

In dieser Näherung erhalten wir schliesslich mit Hilfe der Gleichungen (13), (14), und (15)

$$\boldsymbol{\sigma}^{(1)} = \frac{1}{3\pi^2}\cdot\frac{e^2}{\hbar^2}\left\{k^2\cdot\frac{d\epsilon}{dk}\boldsymbol{\tau} + \frac{\pi^2}{6}(KT)^2\frac{d^2}{d\epsilon^2}\left(k^2\frac{d\epsilon}{dk}\boldsymbol{\tau}\right)\right\}_{\epsilon=\zeta} \tag{58}$$

$$\boldsymbol{\varkappa}^{(1)} = \frac{1}{9}\cdot\frac{K^2}{\hbar^2}T\left\{k^2\frac{d\epsilon}{dk}\boldsymbol{\tau}\right\}_{\epsilon=\zeta} = \frac{\pi^2}{3}\cdot\frac{K^2}{e^2}T\cdot\boldsymbol{\sigma}^{(1)} \tag{59}$$

$$\boldsymbol{S}^{(1)} = -\frac{\pi^2}{3}\cdot\frac{K^2}{e}T\cdot\left\{\left(k^2\frac{d\epsilon}{dk}\boldsymbol{\tau}\right)^{-1}\frac{d}{d\epsilon}\left(k^2\frac{d\epsilon}{dk}\boldsymbol{\tau}\right)\right\}_{\epsilon=\zeta} = -\frac{\pi^2}{3}\cdot\frac{K^2}{e}T\left\{\boldsymbol{\sigma}^{(1)-1}\frac{d}{d\epsilon}\boldsymbol{\sigma}^{(1)}\right\}_{\epsilon=\zeta}. \tag{60}$$

Bei Vernachlässigung höherer Glieder in KT/ζ ist der von einem Gitterfehler bei sehr tiefen Temperaturen erzeugte elektrische Widerstand identisch mit dem entsprechenden zusätzlichen Widerstand, den SÁENZ[7] für das Gebiet oberhalb der

* Damit keine Verwechslung mit den in anderen Abschnitten eingeführten Grössen entsteht, bezeichnen wir die durch den Übergang zur reellen Schreibweise erhaltenen Grössen mit den entsprechenden kleinen Buchstaben.

† Zwischen dem von Saenz eingeführten linearen Integraloperator $\mathfrak{Q}_1$ und dem hier verwandten Operator $\mathfrak{Q}$ besteht folgender Zusammenhang :

$$\mathfrak{Q} = \frac{\partial f}{\partial \epsilon}\,\mathfrak{Q}_1.$$

Debyetemperatur berechnet hat. Dieselbe Gesetzmässigkeit gilt auch für den Wärmewiderstand. Die Matthiessensche Regel für die Additivität der elektrischen bezw. der Wärmewiderstände ist bei einem stark entarteten Elektronengas dann erfüllt, wenn sich die durch Gitterfehler erzeugte Störung der Verteilungsfunktion durch Kugelfunktionen erster Ordnung beschreiben lässt. In einer weiteren Arbeit werden wir zeigen, dass in dieser Näherung auch keine Änderung des elektrischen Widerstandes im Magnetfeld auftritt.

8. DER EINFLUSS DER SYMMETRIE DES GITTERFEHLERS

Bei ausgeprägt anisotrop streuenden Gitterfehlern — Versetzungen, Stapelfehlern und dgl. — lässt sich die gestörte Verteilungsfunktion nicht genügend genau durch Kugelfunktionen erster Ordnung beschreiben. In diesen Fällen müssen für die Winkelabhängigkeit höhere Näherungen verwendet werden. Die Bestimmung der Entwicklungskoeffizienten $\mathfrak{P}^{(a)m}_{n}$ mit Hilfe von Gleichung (42) lässt sich sehr oft vereinfachen, wenn bei der Auswahl der Vergleichsfunktionen die Symmetrie der Übergangswahrscheinlichkeit berücksichtigt wird.

(a) Isotropes Streuzentrum

Bei einem isotropen Streuzentrum ist die Übergangswahrscheinlichkeit für einen Sprung von $\mathfrak{k}$ nach $\mathfrak{k}'$ nur vom Winkel Θ zwischen $\mathfrak{k}$ und $\mathfrak{k}'$ abhängig. Für eine bestimmte Energie ϵ lässt sich dann die Übergangswahrscheinlichkeit durch folgende Entwicklung darstellen

$$w(\mathfrak{k},\mathfrak{k}') = \sum_{\nu=0}^{\infty} \frac{2\nu+1}{2} \cdot w_\nu(\epsilon) P_\nu(\cos\Theta). \tag{61}$$

Bei sphärischen Polarkoordinaten ist

$$\begin{aligned}\cos\Theta &= \cos\vartheta\cdot\cos\vartheta' + \\ &+\sin\vartheta\cdot\sin\vartheta'\cdot\cos(\phi-\phi').\end{aligned} \tag{62}$$

Mit dem Additionstheorem der Kugelfunktionen[20] ergibt sich dann

$$\begin{aligned}w(\mathfrak{k},\mathfrak{k}') = \sum_{\nu=0}^{\infty}\sum_{\mu=-\nu}^{\nu} w_\nu(\epsilon)\Pi^\mu_\nu(\cos\vartheta)\times \\ \times\Pi^\mu_\nu(\cos\vartheta')\cdot e^{i\mu(\phi-\phi')}.\end{aligned} \tag{63}$$

Für das Matrixelement $T^{m\,o}_{n\,l}$ folgt dann

$$T^{m\,o}_{n\,l} = k^2\frac{dk}{d\epsilon}\cdot\frac{1}{\tau_n}\delta_{n,l}\cdot\delta_{m,-o} \tag{64a}$$

mit

$$\frac{1}{\tau_n} = \frac{V}{(2\pi)^2}k^2\frac{dk}{d\epsilon}[w_0(\epsilon)-w_n(\epsilon)]. \tag{64b}$$

Das lineare Gleichungssystem (42) zerfällt also für jedes n und m in voneinander unabhängige Gleichungssysteme. Berücksichtigen wir noch, dass $\mathfrak{U}^m_n(\epsilon)\equiv 0$, wenn nicht $n=1$ und $m=0$ oder ± 1 ist, so lassen sich diese Gleichungssysteme sicher erfüllen, wenn für alle $n>1$ und $|m|>1$ die Entwicklungskoeffizienten $\mathfrak{P}^{(a)m}_{n}$ identisch Null sind.

Wie wir noch sehen werden, sind die Matrixelemente $T^{mo}_{n\,l}$ mit $n\neq l$ für die Abweichungen von der Matthiessenschen Regel und für die Änderung des elektrischen Widerstandes im Magnetfeld entscheidend. In der Näherung eines quasifreien, stark entarteten Elektronengases sind also diese Effekte bei isotropem Streuer nicht zu erwarten.

(b) Gitterfehler mit Symmetrieebene und darauf senkrecht stehender zwei-zähliger Symmetrieachse

Für eine bestimmte Energie ϵ lässt sich bei einem quasifreien Elektronengas die Übergangswahrscheinlichkeit $w(\mathfrak{k},\mathfrak{k}')$ ganz allgemein durch folgende Reihenentwicklung darstellen

$$\begin{aligned}w(\mathfrak{k},\mathfrak{k}') = \sum_{\nu=0}^{\infty}\sum_{\lambda=0}^{\infty}\sum_{\mu=-\nu}^{\nu}\sum_{\omega=-\lambda}^{\lambda} w^{\mu\ \omega}_{\nu\ \lambda}(\epsilon)\times \\ \times\Pi^\mu_\nu(\cos\vartheta)\cdot\Pi^\omega_\lambda(\cos\vartheta')e^{i(\mu\phi+\omega\phi')}.\end{aligned} \tag{65}$$

Für das Matrixelement $T^{mo}_{n\,l}$ ergibt sich dann

$$\begin{aligned}T^{mo}_{n\,l}(\epsilon) = \frac{(-1)^o}{(2\pi)^2}V\cdot k^4\left(\frac{dk}{d\epsilon}\right)^2\times \\ \times\Big[\sum_{\nu=0}^{\infty} w^{-m-o\ 0}_{\nu\ \ 0}(\epsilon)\int_{-1}^{1}\Pi^m_n\Pi^o_l\Pi^{m+o}_\nu dx - \\ -w^{-m-o}_{n\ \ l}(\epsilon)\Big].\end{aligned} \tag{66}$$

Besitzt das Störpotential eine zweizählige Symmetrieachse, die mit der Rotationsachse eines

sphärischen Polarkoordinatensystems zusammenfällt, so ist $w(\mathfrak{k}, \mathfrak{k}')$ invariant gegenüber einer Drehung des Koordinatensystems um den Winkel $\phi_0 = \pi$. Alle Entwicklungskoeffizienten in (66) sind Null, wenn nicht

$$\mu+\omega = 2\rho \qquad \rho = 0, 1, 2 \ldots.$$

Das Matrixelement $T^{mo}_{n\,l}$ ist also nur dann von Null verschieden, wenn m und o entweder beide gerade oder ungerade sind. Ist zudem noch die x-y Ebene eine Symmetrieebene, d.h. ist $w(\mathfrak{k}, \mathfrak{k}')$ invariant gegenüber einer gleichzeitigen Vertauschung von k_z und k_z' gegenüber $-k_z$ und $-k_z'$, so sind die von Null verschiedenen Matrixelemente bestimmt durch

$$\text{d.h.} \quad \begin{aligned} \nu+\lambda+\mu+\omega &= 2\rho' \qquad \rho' = 0, 1, 2, 3, \ldots \\ \nu+\lambda &= 2\rho'' \qquad \rho'' = 0, 1, 2, 3, \ldots . \end{aligned}$$

Wie man aus der allgemeinen Darstellung entnehmen kann, überträgt sich diese Eigenschaft von $w^{\mu\omega}_{\nu\,\lambda}$ auf das Matrixelement $T^{mo}_{n\,l}$:

$$T^{m\,o}_{n\,l} = 0, \text{ wenn}$$

$$\begin{aligned} n+l &= 2r+1 \qquad r = 0, 1, 2, \ldots \\ m+o &= 2r'+1 \qquad r' = 0, 1, 2, \ldots . \end{aligned} \tag{67}$$

Das lineare Gleichungssystem (42) zerfällt dadurch in 4 voneinander unabhängige Gleichungssysteme, je nachdem m und n gerade oder ungerade ist. Berücksichtigen wir noch, dass bei einem quasifreien Elektronengas $\mathfrak{U}^m_n(\epsilon)=0$ für $n \neq 1$ ist, so brauchen wir nur das lineare Gleichungssystem für die Koeffizienten mit ungeradem n zu lösen, während die Koeffizienten mit geradem n identisch Null sind.

(c) Rotationssymmetrischer Gitterbaufehler

Bei einem rotationssymmetrischen Streupotential kann die Übergangswahrscheinlichkeit $w(\mathfrak{k}, \mathfrak{k}')$ nur von der Differenz $(\phi-\phi')$ abhängen. Die Entwicklungskoeffizienten in (65) müssen deshalb folgende Forderung erfüllen :

$$w^{\mu\ \omega}_{\nu\ \lambda}(\epsilon) = w^{\mu-\mu}_{\nu\ \lambda}(\epsilon) \cdot \delta_{\mu,-\omega} \tag{68}$$

Diese Eigenschaft überträgt sich auch auf die Matrixelemente $T^{mo}_{n\,l}$

$$T^{m\,o}_{n\,l} = T^{m-m}_{n\ \ l} \cdot \delta_{m,-o}. \tag{69}$$

Besitzt das rotionssymmetrische Potential ***keinen*** Windungssinn, so ist $w(\mathfrak{k}, \mathfrak{k}')$ nur vom Betrage $|\phi-\phi'|$ abhängig. Es muss dann sein

$$w^{\mu-\mu}_{\nu\ \lambda} = w^{-\mu\ \mu}_{\nu\ \lambda}. \tag{70}$$

Bei rotationssymmetrischen Problemen ohne Windungssinn ist das Matrixelement $T^{mo}_{n\,l}$ invariant gegenüber einer Vertauschung der oberen Indizes. Berücksichtigen wir ausserdem (17), so ergibt sich

$$T^{m-m}_{n\ \ l} \equiv T^{-m\,m}_{n\,l} \equiv T^{m-m}_{l\ \ n}. \tag{71}$$

Durch diese besondere Eigenschaft des Matrixelementes $T^{mo}_{n\,l}$ zerfällt das lineare Gleichungssystem (42) in $2q+1$ voneinander unabhängige Gleichungssysteme, die zu verschiedenen oberen Indizes m gehören. Für die Berechnung der Leitfähigkeitstensoren brauchen dabei nur die Gleichungssysteme mit $m = 0$ und $m = \pm 1$ gelöst werden.

(d) Gitterfehler mit kubischer Symmetrie

Bei einem Gitterfehler mit kubischer Symmetrie ist die Übergangswahrscheinlichkeit $w(\mathfrak{k}, \mathfrak{k}')$ invariant gegen folgende Symmetrieoperationen :

(1) Drehung des Koordinatensystems um den Winkel $\alpha = \pi/2$ mit k_z als Drehachse
(2) Spiegelung an der Ebene $k_z = 0$
(3) Drehung des Koordinatensystems um den Winkel $\beta = \pi/2$ mit k_y als Drehachse.

Die ersten beiden Symmetriebedingungen werden durch folgende Entwicklung erfüllt*:

$$w(\mathfrak{k}, \mathfrak{k}') = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \sum_{\lambda} \sum_{\omega} w^{\mu\ 4\omega-\mu}_{\nu\ 2\lambda+\nu} \times \tag{72}$$

$$\times \Pi^{\mu}_{\nu}(\cos\vartheta) e^{i\mu\phi} \cdot \Pi^{4\ \omega-\mu}_{2\lambda+\nu}(\cos\vartheta') \cdot e^{i(4\omega-\mu)\phi'}.$$

Um den Einfluss der dritten Symmetriebedingung auf die Übergangswahrscheinlichkeit zu untersuchen, entwickeln wir die in (72) vorkommenden Kugelflächenfunktionen nach Kugelflächenfunktionen, die die k_y-Achse als Polarachse

* Die Bezeichnung ist so gewählt, dass mit

$$w^{\mu\ 4\omega-\mu}_{\nu\ 2\lambda+\nu} = w^{\mu-\mu}_{\nu\ \nu}\, \delta_{\lambda,0} \cdot \delta_{\omega,0}$$

isotrope Verhältnisse entstehen.

haben. Hierbei berücksichtigen wir, dass bei einer Drehung $\mathfrak{D}$ des Koordinatensystems um die Eulerschen Winkel α, β, und γ sich die Kugelflächenfunktionen n-ter Ordnung in eine Linearkombination von Kugelflächenfunktionen derselben Ordnung, aber verschiedener oberer Indizes aufspalten[21]*

$$\Pi_n^m(\cos\vartheta)\cdot e^{im\phi}$$

$$=\sum_{m'=-n}^{n}\mathfrak{D}^{(n)}\{\alpha,\beta,\gamma\}_{mm'}\,\Pi_n^{m'}(\cos\Theta)\cdot e^{im'\Phi}. \qquad (73)$$

Setzen wir diese Entwicklung mit $\alpha=\gamma=0$ und $\beta=\pi/2$ in (72) ein und berücksichtigen, dass wegen der dritten Symmetriebedingung dieselbe Entwicklung um k_y als Polarachse durchgeführt werden kann, so ergibt sich folgender Zusammenhang zwischen den Entwicklungskoeffizienten derselben Ordnung

$$\sum_{\mu}\sum_{\omega} w_{\nu\,2\lambda+\nu}^{\mu\;4\omega-\mu}\cdot\mathfrak{D}^{(\nu)}\left(0;\frac{\pi}{2};0\right)_{\mu,\mu'}\times$$
$$\times\mathfrak{D}^{(2\lambda+\nu)}\left(0;\frac{\pi}{2};0\right)_{4\omega-\mu,\,4\omega'-\mu'}=w_{\nu\;2\lambda+\nu}^{\mu'\;4\;\omega'-\mu'}. \qquad (74)$$

Die von Null verschiedenen Matrixelemente T_{nl}^{mo} sind gegeben durch

$$T_{n\;2l+n}^{m\;4o-m}=\frac{(-1)^{4o-m}}{(2\pi)^2}Vk^4\left(\frac{dk}{d\epsilon}\right)^2\times$$
$$\times\left[\sum_{\nu=0}^{\infty}w_{2\;\nu\;0}^{-4o\;0}\int_{-1}^{1}\Pi_n^m(x)\cdot\Pi_{2l+n}^{4o-m}(x)\Pi_{2\nu}^{4o}(x)\,dx-\right.$$
$$\left.-w_{n\;\;2l+n}^{-m\;\;-4o+m}\right]. \qquad (75)$$

Gleichung (74) hat noch zur Folge, dass die Matrixelemente $T_{n\;2l+n}^{m\;4o-m}$ mit gleichen unteren Index nicht linear unabhängig sind.

* Man beachte jedoch, dass sich die von Wigner verwendeten normierten Kugelfunktionen von den unsrigen um den Faktor i^m unterscheiden.

9. DIE TRANSPORTERSCHEINUNGEN BEI GLEICHZEITIGER ANWESENHEIT VON ISOTROP UND ANISOTROP STREUENDEN GITTERFEHLERN

In diesem Abschnitt soll der Fall behandelt werden, dass neben anisotrop streuenden Gitterfehlern auch noch isotrop streuende Gitterfehler vorhanden sind. Die einzelnen Fehlstellen sollen so weit voneinander entfernt sein, dass die Vielfachstreuung der Elektronen vernachlässigt werden kann. Der gesamte Kristall lässt sich dann in einzelne Volumgebiete einteilen, in denen nur Streuprozesse durch den in diesem Gebiet vorkommenden Gitterfehler verursacht werden. Eine genaue Betrachtung zeigt, dass sich in dieser Näherung die Übergangswahrscheinlichkeiten einfach addieren. Ist ein Teil der anisotropen Fehlstellen so angeordnet, dass ihre ausgezeichneten Richtungen im Gitter isotrop verteilt sind, so hat ihr Beitrag zum Matrixelement T_{nl}^{mo}, wie anschaulich verständlich ist und sich (unter Vernachlässigung der Vielfachstreuung) mathematisch zeigen lässt, dieselben Eigenschaften wie bei isotropen Streuern. Wir denken uns im folgenden diesen Anteil zu den isotropen Streuern hinzugerechnet.

Ist Z_{I} die Zahl der ausgerichteten anisotropen und Z_{II} die Zahl der isotropen Gitterfehlstellen, so erhalten wir statt (9) folgende Integralgleichung

$$Z_{\mathrm{I}}\cdot\mathfrak{L}_{\mathrm{I}}\{\mathfrak{X}(\mathfrak{k})\}+Z_{\mathrm{II}}\cdot\mathfrak{L}_{\mathrm{II}}\{\mathfrak{X}(\mathfrak{k})\}$$
$$=\frac{1}{\hbar}\cdot\frac{\partial f_0}{\partial\epsilon}\,\mathrm{grad}_{\mathfrak{k}}\epsilon, \qquad (76)$$

dabei ist $\mathfrak{L}_{\mathrm{I}}$ der Integraloperator für *eine* anisotrope und $\mathfrak{L}_{\mathrm{II}}$ der Integraloperator für *eine* isotrope Gitterfehlstelle.

Die Lösung der Integralgleichung (76) kann wieder mit dem Variationsverfahren erfolgen. Der einzige Unterschied zum seither beschriebenen Verfahren besteht darin, dass sich das Matrixelement T_{nl}^{mo} additiv aus zwei Grössen zusammensetzt

$$T_{nl}^{m\,o}=Z_{\mathrm{I}}\cdot T_{nl}^{m\,o}+Z_{\mathrm{II}}\cdot T_{\mathrm{II}\,nl}^{\;\;m\,o}, \qquad (77)$$

wobei das von der isotropen Fehlstelle herrührende Matrixelement $T_{\mathrm{II}\,nl}^{\;\;mo}$ wegen (64a) die spezielle Eigenschaft

$$T_{\mathrm{II}}{}_{nl}^{m\,o}=k^2\frac{dk}{d\epsilon}\tau_n^{-1}\cdot\delta_{n,\,l}\cdot\delta_{o,\,-m}\quad,$$

besitzt. Die zusätzlichen isotropen Streuer haben also bloss einen Einfluss auf das Matrixelement $T^{m-m}_{n\ n}$.

10. DAS TRANSPORTPROBLEM OBERHALB DER DEBYETEMPERATUR Θ_D. ABWEICHUNGEN VON DER MATTHIESSENSCHEN REGEL

Mit den Ergebnissen aus Abschnitt 9 können wir das Transportproblem oberhalb der Debyetemperatur lösen. Die Streuung der Elektronen an den Gitterschwingungen und an den Gitterfehlern lässt sich zwar nicht in der beschriebenen Weise trennen; doch zeigt eine nähere Untersuchung,[22] das sich die entsprechenden Übergangswahrscheinlichkeiten addieren. In der Näherung quasifreier Elektronen ist bei Vernachlässigung von Umklappprozessen die Übergangswahrscheinlichkeit w_{II} ($\mathfrak{k}$, $\mathfrak{k}'$) für die thermische Streuung nur vom Winkel zwischen $\mathfrak{k}$ und $\mathfrak{k}'$ abhängig; ausserdem erfolgt die thermische Streuung oberhalb der Debyetemperatur elastisch. Alle hier gemachten Voraussetzungen sind erfüllt, so dass wir mit dem beschriebenen Variationsverfahren die Änderung der Leitfähigkeitsgrössen durch anisotrope Fehlstellen oberhalb der Debyetemperatur berechnen können. Gegenüber dem Iterationsverfahren von SÁENZ,[7] bei dem die durch die Gitterfehlstellen bedingte Übergangswahrscheinlichkeit als kleine Störung der thermischen Streuung aufgefasst wird, hat das Variationsverfahren den Vorteil, dass die durch die thermische Streuung geänderte Elektronenverteilung bloss in derselben Genauigkeit bestimmt zu werden braucht, wie es für die anisotrope Streuung gewünscht wird. Beim Iterationsverfahren müssen dagegen die durch die thermische Streuung bestimmten Grössen exakt bekannt sein. Ein weiterer Vorteil des Variationsverfahren besteht darin, dass mit ihm sowohl die Transportgrössen bei tiefer als auch bei hoher Temperatur bestimmt werden können. Es ist also möglich, mit diesem Verfahren die Abweichungen von der Matthiessenschen Regel über die Additivität des thermischen und elektrischen Widerstandes zu untersuchen, worauf wir zum Abschluss noch kurz eingehen werden.

Zuvor wollen wir jedoch die Transportgrössen oberhalb der Debyetemperatur in der ersten Näherung für die Winkelabhängigkeit angeben und zeigen, dass — wie schon im Abschnitt 7 erwähnt wurde — bei einem stark entarteten Elektronengas die zusätzlichen Widerstände mit den entsprechenden Restwiderständen bei tiefen Temperaturen identisch sind. Die Wirkung von anisotrop streuenden Gitterfehlstellen haben wir in Abschnitt 7 durch den Tensor $\boldsymbol{\tau}$ beschrieben. Sind ausserdem noch isotrop streuende Fehlstellen vorhanden, so tritt zu dem Tensor $\boldsymbol{\tau}^{-1}$ noch der Kugeltensor $\tau_1^{-1}\mathbf{1}$ hinzu; τ_1 ist dabei die Relaxationszeit für die gewöhnliche Elektron-Phonon Wechselwirkung.

Wenn wir bis zur zweiten Potenz im Entartungsparameter KT/ζ gehen, ist der elektrische Widerstand gegeben durch*

$$\boldsymbol{\rho}^{(1)} = 3\pi^2\frac{\hbar^2}{e^2}\left\{\frac{\boldsymbol{\tau}^{-1}+\tau_1^{-1}\mathbf{1}}{k^2\,d\epsilon/dk}\left[\mathbf{1}-\frac{\pi^2}{6}(KT)^2\times\right.\right.$$
$$\times\frac{d^2}{d\epsilon^2}\left(k^2\frac{d\epsilon}{dk}(\boldsymbol{\tau}^{-1}+\tau_1^{-1}\mathbf{1})^{-1}\right)\times \tag{78}$$
$$\left.\left.\times\frac{(\boldsymbol{\tau}^{-1}+\tau_1^{-1}\mathbf{1})}{k^2\,d\epsilon/dk}\right]\right\}_{\epsilon=\zeta}.$$

Im strengen Sinne ist also die Matthiessensche Regel über die Additivität der elektrischen Widerstände nicht erfüllt. Bei einem stark entarteten Elektronengas ist jedoch die Korrektur vernachlässigbar.

In derselben Weise kann man zeigen, dass Abweichungen von der Additivität der Wärmewiderstände erst dann auftreten, wenn man die Integration über ε bis zur vierten Potenz im Entartungsparameter durchführt.

$$(\boldsymbol{\varkappa}^{(1)})^{-1} = 9\frac{\hbar^2}{K}\cdot\frac{1}{T}\left\{\frac{\boldsymbol{\tau}^{-1}+\tau_1^{-1}\mathbf{1}}{k^2\,d\epsilon/dk}\right\}_{\epsilon=\zeta}$$
$$= (\boldsymbol{\varkappa}^{(1)}{}_{\text{Rest}})^{-1}+(\boldsymbol{\varkappa}^{(2)}{}_{\text{thermisch}})^{-1} \tag{79}$$

Da oberhalb der Debyetemperatur $\boldsymbol{\tau}^{-1} \ll \tau_1^{-1}\mathbf{1}$ ist, gilt auch für die Änderung der Wärmeleitfähigkeit und der elektrischen Leitfähigkeit das Wiedemann-Franzsche Gesetz

$$\boldsymbol{\varkappa}^{(1)} - \boldsymbol{\varkappa}^{(1)}{}_{\text{thermisch}} = \frac{\pi}{3}\cdot\frac{K^2T}{e^2}(\boldsymbol{\sigma}^{(1)} - \boldsymbol{\sigma}^{(1)}{}_{\text{thermisch}}). \tag{80}$$

Es ist dabei zu beachten, dass bei tiefer Temperatur die Restwärmeleitfähigkeit proportional zur absoluten Temperatur ist, während die Änderung der Wärmeleitfähigkeit oberhalb der Debyetemperatur mit T^{-1} abnimmt.

Für die Thermokraft erhalten wir schliesslich

$$\boldsymbol{S}^{(1)} = -\frac{\pi^2}{3}\cdot\frac{K^2}{e}T\left\{\left(k^2\frac{dk}{d\epsilon}\right)^{-1}(\boldsymbol{\tau}^{-1}+\tau_1^{-1}\mathbf{1})\times\right.$$
$$\left.\times\frac{d}{d\epsilon}\left[k^2\frac{dk}{d\epsilon}(\boldsymbol{\tau}^{-1}+\tau_1^{-1}\mathbf{1})^{-1}\right]\right\}_{\epsilon=\zeta}, \tag{81}$$

* Wir denken uns $Z_I = Z_{II} = 1$ gesetzt. Ist Z_I und $Z_{II} \neq 1$, dann müssen wir $\boldsymbol{\tau}^{-1}$ durch $Z\,\boldsymbol{\tau}^{-1}$ und τ_1^{-1} durch $Z_{II}\,.\,\tau_1^{-1}$ ersetzen.

VOL. 4 1958

bezw. für ihre Änderung

$$\boldsymbol{S}^{(1)}-\boldsymbol{S}^{(1)}_{\text{thermisch}}=+\frac{\pi^2}{3}\cdot\frac{K^2}{e}T\left\{\frac{d}{d\epsilon}(\tau_1\tau^{-1})\right\}_{\epsilon=\zeta}. \tag{82}$$

Die Änderung der Thermokraft oberhalb der Debyetemperatur ist proportional zur Fehlstellenkonzentration und temperaturunabhängig, während sie bei sehr tiefen Temperaturen proportional zur absoluten Temperatur wird.

Wie schon im Abschnitt 7 erwähnt wurde, treten Abweichungen von der Matthiessenschen Regel erst bei den höheren Näherungen für die Winkelabhängigkeit auf, die nicht vom Entartungsparameter abhängig sind, also auch bei einem stark entarteten Elektronengas nicht verschwinden. Da die höheren Näherungen im wesentlichen von der Symmetrie des Streuers abhängen, wird auf eine allgemeine Berechnung dieser Abweichungen verzichtet. Die wesentlichen Gesichtspunkte lassen sich am Beispiel eines rotationssymmetrischen Potentials erläutern, das noch eine auf der Rotationsachse senkrecht stehende Symmetrieebene hat.

Für die Transportgrössen senkrecht und parallel zur Rotationsachse ergibt sich, wenn wir bis zu Kugelfunktionen dritter Ordnung gehen und uns auf ein stark entartetes Elektronengas beschränken (Berücksichtigung der ersten im Entartungsparameter nicht verschwindenden Glieder)*

$T \ll \Theta_D$:

$$\rho_{\perp}^{(3)}=3\pi^2\frac{\hbar^2}{e^2}\cdot\frac{1}{k_F{}^4}\left(T^{1-1}_{1\ 1}-\frac{T^{1-1}_{1\ 3}T^{1-1}_{3\ 1}}{T^{1-1}_{3\ 3}}\right)_{\epsilon=\zeta} \tag{83a}$$

$$\rho_{\parallel}^{(3)}=3\pi^2\frac{\hbar^2}{e^2}\cdot\frac{1}{k_F{}^4}\left(T^{00}_{11}-\frac{T^{00}_{13}T^{00}_{31}}{T^{00}_{33}}\right)_{\epsilon=\zeta}; \tag{83b}$$

$T > \Theta_D$:

$$\rho_{\perp}^{(3)}=3\pi^2\frac{\hbar^2}{e^2}\cdot\frac{1}{k_F{}^4}\times \times\left(T^{1-1}_{1\ 1}+k^2\frac{dk}{d\epsilon}\tau_1^{-1}-\frac{T^{1-1}_{13}T^{1-1}_{3\ 1}}{T^{1-1}_{3\ 3}+k^2\,dk/d\epsilon\,\tau^{-1}_3}\right)_{\epsilon=\zeta} \tag{84a}$$

$$\rho_{\parallel}^{(3)}=3\pi^2\frac{\hbar^2}{e^2}\cdot\frac{1}{k_F{}^4}\times \times\left(T^{00}_{11}+k^2\frac{dk}{d\epsilon}\tau_1^{-1}-\frac{T^{00}_{13}T^{00}_{31}}{T^{00}_{33}+k^2\,dk/d\epsilon\,\tau_3^{-1}}\right)_{\epsilon=\zeta}; \tag{84b}$$

$$\kappa^{(3)}=\frac{\pi^2}{3}\cdot\frac{K^2}{e^2}T\frac{1}{\rho^{(3)}}. \tag{85}$$

Die Abweichungen von der Matthiessenschen Regel sind demnach

$$(\Delta\rho_{\perp})_{\text{Matt}}=3\pi^2\frac{\hbar^2}{e^2}\left\{\frac{1}{k^4}T^{1-1}_{1\ 3}T^{1-1}_{3\ 1}\left(\frac{1}{T^{1-1}_{3\ 3}}-\frac{1}{T^{1-1}_{3\ 3}+k^2(dk/d\epsilon)\tau_3{}^{-1}_3}\right)\right\}_{\epsilon=\zeta}$$

$$\approx 3\pi^2\frac{\hbar^2}{e^2}\left\{\frac{1}{k^4}\cdot\frac{T^{1-1}_{1\ 3}\cdot T^{1-1}_{3\ 1}}{T^{1-1}_{3\ 3}}\right\}_{\epsilon=\zeta}. \tag{86a}$$

$$(\Delta\rho_{\parallel})_{\text{Matt}}=3\pi^2\frac{\hbar^2}{e^2}\left\{\frac{1}{k^4}T^{00}_{13}\cdot T^{00}_{31}\left(\frac{1}{T^{00}_{33}}-\frac{1}{T^{00}_{33}+k^2(dk/d\epsilon)\tau_3^{-1}}\right)\right\}_{\epsilon=\zeta}$$

$$\approx 3\pi^2\frac{\hbar^2}{e^2}\left\{\frac{1}{k^4}\cdot\frac{T^{00}_{13}\cdot T^{00}_{31}}{T^{00}_{33}}\right\}_{\epsilon=\zeta} \tag{86b}$$

Bei der letzten Gleichung wurde schon benutzt, dass $\tau_3^{-1}k^2(dk/d\epsilon) \gg T^{1-1}_{3\ 3}$ bezw. T^{00}_{33} ist.

Die Abweichungen von der Matthiessenschen Regel sind also oberhalb der Debyetemperatur nahezu temperaturunabhängig und entsprechen gerade den Widerstandsänderungen, die durch die höheren Näherungen bedingt sind. Dieses Verhalten erwarten wir ganz allgemein, da oberhalb der Debyetemperatur $k^2(dk/d\epsilon)\tau_n^{-1} \gg T^{m-m}_{n\ n}$. Aus dem Beispiel ist ferner zu entnehmen, dass die Abweichungen von der Matthiessenschen Regel durch die Matrixelemente T^{mo}_{1i} bedingt sind,

* Bei den Matrixelementen T^{mo}_{ln} haben wir den unteren Index I weggelassen.

die nur bei anisotropen Gitterfehler von Null verschieden sind.

Die Verfasser danken Herrn Professor U. DEHLINGER für sein förderndes Interesse und der Deutschen Forschungsgemeinschaft für die Unterstützung der Arbeit.

LITERATUR

1. MACKENZIE J. K. und SONDHEIMER E. H. *Phys. Rev.* **77,** 264 (1950).
2. LANDAUER R. *Phys. Rev.* **82,** 520 (1951).
3. DEXTER D. L. *Phys. Rev.* **86,** 770 (1952).
4. HIRONE T. und ADACHI K. *Sci. Rep. Res. Insts. Tôhoku Univ.* A **3,** 454 (1951).
5. HUNTER S. C. und. NABARRO F. R. N. *Proc. Roy. Soc.* A **220,** 542 (1953).
6. SEEGER A. und STEHLE H. *Z. Phys.* **146,** 242 (1956).
7. SÁENZ A. W. *Phys. Rev.* **91,** 1142 (1953).
8. NORDHEIM L. *Ann. Phys., Lpz.* **9,** 607 (1931).
9. MOTT N. F. *Proc. Camb. Phil. Soc.* **32,** 281 (1936).
10. SEEGER A. *The International Conference on Electron Transport in Metals and Solids.* Ottawa, September (1956); *Canad. J. Phys.* **34,** 1219 (1956).
11. VAN BUEREN H. G. *Colloquium on Plastic Flow and Deformation of Solids, Madrid* 1955 S.79. Springer, Berlin, Göttingen, Heidelberg (1956).
12. VAN BUEREN H. G. *Dissertation Leiden* (1956).
13. KOHLER M. *Z. Phys.* **125,** 679 (1949).
14. ENSKOG D. *Dissertation Uppsala* (1917).
15. NORDHEIM L. *Ann. Phys., Lpz.* **9,** 607 (1931).
16. KOHLER M. *Ann. Phys., Lpz.* **40,** 601 (1941).
17. KOHLER M. *Z. Phys.* **124,** 772 (1948).
18. KOWALEWSKI G. *Determinantentheorie.* Berlin, Leipzig (1925).
19. MOTT N. F. und JONES H. *The Theory of the Properties of Metals and Alloys.* Oxford (1936).
20. SOMMERFELD A. *Partielle Differentialgleichungen der Physik* (2. Auflage) Wiesbaden (1947).
21. WIGNER E. *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren.* Braunschweig (1931).
22. JONES H. *The Theory of Electrical and Thermal Conductivity in Metals, Handbuch der Physik XIX.* Springer (1956).

VOL.
4
1958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 177–181.

COMPARISON OF ELECTRON WITH NEUTRON IRRADIATION OF α-BRASS*

A. C. DAMASK†

Pitman-Dunn Laboratories, Frankford Arsenal, Philadelphia, Pennsylvania

(*Received* 26 *June* 1957)

Abstract—The decrease in electrical resistivity of α-brass which arises from increased short-range order is used as a measure of migration of excess vacancies during neutron and electron irradiation. It is shown that irradiation by 2 MeV electrons from a Van de Graaff generator will decrease the resistivity at the same rate as irradiation in the *BNL* reactor. However, theoretical estimates indicate that about fourteen times more vacancies are produced by the electron than by the neutron irradiation. The results demonstrate, however, that electron irradiation, although less efficient, is a feasible method for enhancing order in an alloy at temperatures well below those at which ordering can be achieved by thermal annealing alone.

INTRODUCTION

THE diffusion of atoms in a substitutional solid solution in general obeys an Arrhenius equation

$$\tau = \tau_0 \, e^{H/RT} \qquad (1)$$

where τ is the mean jump time of an atom, τ_0 is a constant of the material, H is the activation energy of the process, R the gas constant, and T the absolute temperature. The activation energy H is the energy barrier which an atom must cross in order to change its location in the lattice to an adjacent site. In a solid solution which uses the vacancy mechanism of diffusion, an atom can move only when there exists a vacant lattice site in an adjacent position. The total activation energy H for atomic migration must therefore be the sum of two activation energies: the energy H_f required to form a vacancy, and the energy H_J required for an atom to jump into the vacancy. A recent review of existing data by LI and NOWICK[1] indicates that H_J/H is about 0·6. That is, 60 per cent of the activation energy for atomic mobility is required for the jump energy, H_J, and 40 per cent for the energy to form a vacancy, H_f. The addition of vacancies in excess of the thermal equilibrium number can therefore reduce H up to 40 per cent, depending upon the number of excess vacancies introduced. Since H appears in the exponent of equation (1), a substantial reduction can change the mean jump time τ by orders of magnitude.‡

When a solid is bombarded with energetic particles, one of the defects created is the vacancy-interstitial pair,[2] i.e. an atom is knocked out of its place in the lattice into an interstitial position and leaves behind a vacant lattice site. This vacancy can then migrate with an activation energy H_J. The introduction of excess vacancies by irradiation will enhance diffusion in an alloy, and can assist the alloy in achieving a lower free-energy state at a temperature where under normal conditions all atoms would be "frozen" in position.

α-Brass has proved to be a useful alloy for studying radiation effects. The atoms do not have to diffuse very far to cause changes in the short-range order, and it has been shown that small changes in short-range order cause measurable changes in the electrical resistivity. Previous studies

* Supported in part by the U.S. Atomic Energy Commission.

† Guest Scientist at Brookhaven National Laboratory, Upton, New York.

‡ LI and NOWICK also suggest that what are believed to be vacancies created in copper by irradiation are really some faster-moving defect such as divacancies with an energy ratio of $H_J/H = 0{\cdot}35$. This would result in an even greater enhancement of atomic mobility.

of this alloy have also shown that neutron irradiation enhances short-range order and hence lowers the electrical resistivity.[3,4] The disadvantage of neutron irradiation for alloy studies is that the thermal spike associated with the primary knock-on causes damage other than the simple creation of vacancy-interstitial pairs. This other damage can be disruptive if the primary desire is simply to allow the alloy to achieve a lower energy state by the introduction of excess vacancies. Current theory indicates that electron bombardment of metals produces very little, if any, damage other than the creation of vacancy-interstitial pairs. The only previous electron irradiations of an alloy have been carried out on Cu_3Au, and it has been shown that when a partially ordered alloy is irradiated and then raised to the proper temperature for the migration of the vacancies the resistivity decreases slightly.[5–7] This indicates some enhancement of ordering. A point of interest which started the present investigation is whether irradiation with an average electron accelerator for a reasonable length of time can produce the same excess vacancy effects as neutron irradiation. A positive result would establish electron irradiation as a useful tool for studying the low-temperature equilibrium state of alloys.

PROCEDURE

Brass wires (70 per cent Cu, 30 per cent Zn) were drawn to 0·013 in. and annealed at 400°C to remove the cold work. Two potential probe leads were hard soldered on each wire. The wires were then annealed at 210°C for 45 min to establish equilibrium short-range order at that temperature (this procedure is described in reference 3). The electrical resistivities were measured at liquid-nitrogen temperature by the potential-drop method. All resistivities reported in this paper were measured at liquid-nitrogen temperature because it has been shown that resistivity changes due to short-range order changes in α-brass are larger at lower temperatures.[3] The wire for neutron irradiation was about 2 in. long and was sealed in an evacuated quartz tube for the irradiations. It was irradiated in the BNL reactor in a flux of approximately 5×10^{12} neutrons/cm² sec. Of this total flux about 20 to 25 per cent is fast neutrons (the fast neutrons cause most irradiation effects in solids). The fast flux can then be said to be about 10^{12} neutrons/cm² sec. The temperature of the hole used for the irradiation was about 50°C.

The wire used for the electron irradiation was about 1 in. long and was coiled into a small spiral about 4 mm in diameter. An air stream was directed on to the wire to prevent heating by the irradiation. An attempt was made to keep the temperature below 50°C during irradiation, but there is some uncertainty as to the exact temperature. The irradiation was performed with 2 MeV electrons from a Van de Graaff generator. The direction of bombardment was normal to the plane of the spiral. The current was 8 μA and the beam diameter was approximately 0·5 cm. The flux was therefore about $2{\cdot}6\times10^{14}$ electrons/cm² sec. The resistivities were measured at liquid-nitrogen temperature between irradiations. The defects migrate so fast at 50°C that by the time the measurements are made the defects have completely annealed. The resistivity measurements therefore do not detect the defects themselves, but only their effect on the alloy after they have annealed.

RESULTS

The results are shown in Fig. 1, where the broken curve represents neutron and the solid curve electron irradiation. The similarity of the decay rates is remarkable. The neutron curve was extended for 12 hr longer (not shown), but no further decrease was observed. The last point on the electron-irradiation curve seems to indicate that the temperature was about 20°C too high

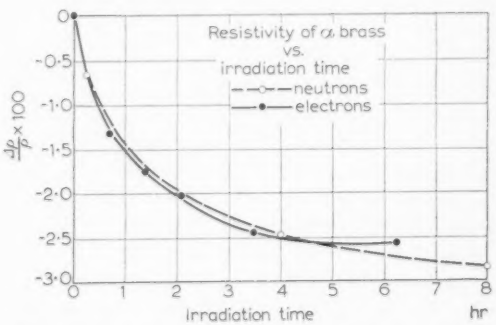


FIG. 1. Resistivity of α-brass vs. irradiation time.

during that one irradiation. When both specimens were annealed at 210°C for 45 min after the completion of the irradiations, they returned to their initial resistivity values.

The solid circles of Fig. 2 show the thermal-equilibrium liquid-nitrogen resistivities of α-brass at various temperatures reported in reference 3.

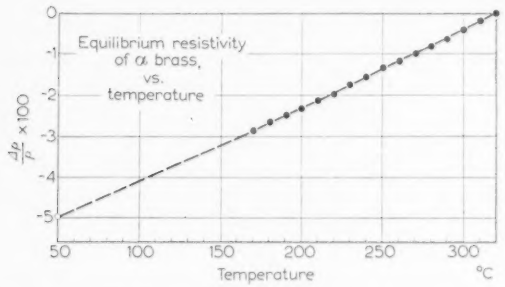


FIG. 2. Equilibrium resistivity of α-brass vs. temperature.

These were obtained by annealing the brass for a suitable length of time at each temperature, then quenching and measuring in liquid nitrogen. It was shown that these changes in resistivity arise from changes in short-range order. The lowest value of the resistivity in the neutron-irradiation curve of Fig. 1 is shown by the open circle at 50°C in Fig. 2. If any long-range order were being induced in the brass, a significant deviation from the short-range order curve might be expected. Since the new point is essentially an extrapolation (dashed portion of Fig. 2) of the short-range order curve, it appears that the critical temperature for long-range order in α-brass must lie below 50°C. The last point on the electron-irradiation curve in Fig. 1 decreased about 0·3 per cent less than the lowest point of the neutron-irradiation curve. The decrease attained by this last electron point corresponds to about 70°C on Fig. 2. This is probably the temperature shift in the specimen mentioned in the preceding paragraph.

The ten-day neutron irradiation reported in reference 3 reduced the resistivity 1·7 per cent below the equilibrium value at 210°C. The 8 hr neutron irradiation shown in Fig. 1 reduced the resistivity 2·86 per cent below the equilibrium value at 210°C. It is not known if further irradiation between 8 hr and 10 days starts to destroy the short-range order produced, or if the dislocation density, and hence trapping of point defects, was different in the specimens. The survey experiments required to determine whether neutrons can destroy short-range order by an irradiation of one sample for times between 8 hr and 10 days take several months to complete because of the radioactivity of the sample. The flux of the *BNL* reactor is being increased in steps over a one-year period. Because of the lack of an accurate fast flux-monitoring system, a correlation of time-dependent effects before and after a flux change cannot be made, so at this time the survey experiments cannot be done.

DISCUSSION

Since the time for decay of the resistivity of the brass is essentially the same for both the neutron and the electron irradiations, it is of interest to compare the particle efficiency. Only crude estimates can be made owing to the lack of precise knowledge of some of the theoretical numbers required.

Because of the size of the wire, the 2 MeV electrons which pass through the center of the wire may lose a small fraction of their energy. The review article by SEITZ and KOEHLER[2] has a convenient plot of the cross-section for displacement of a copper atom as a function of electron energy. For 2 MeV electrons this cross-section is 58 barns and for 1·75 MeV electrons it is 55 barns. For this calculation 57 barns will be used. The electron cross-section varies with Z^2 and inversely with E_d, where Z is the atomic number and E_d the energy to displace an atom into an interstitial position. The Z correction in the cross-section of zinc is negligible, and E_d should not vary very much from that of copper because the atoms are of essentially the same size. Thus, the 57 barn value can be used for the alloy without any significant error. The cross-section for the displacement of a copper atom by 1 MeV neutrons is about 3 barns and that for zinc is 4 barns. Since there is 70 per cent copper in the alloy, if the cross-section for copper (3 barns) is used, the error is not significant for the purposes of this paper. The ratio of the number of primary copper atoms displaced by the incident particle for the electron case to that of the neutron case is

$$\frac{N_e}{N_n} = \frac{\phi_e \sigma_e}{\phi_n \sigma_n} \qquad (2)$$

where ϕ_e and ϕ_n are the fluxes of electrons and neutrons respectively, and σ_e and σ_n are the cross-sections for the electrons and neutrons respectively. For the numbers given, the value of N_e/N_n is about 4900. The probability of an electron displacing two copper atoms or of a copper atom displaced by an electron having enough energy to displace a second is small, so for a first approximation an electron which collides properly can be said to create only one vacancy-interstitial pair. The production of vacancy-interstitial pairs by the primary knock-on after collision with a neutron is less well known. Recent calculations give the value of about 350 vacancy-interstitial pairs.[8] Therefore in this experimental comparison the electron irradiation creates about fourteen times the number of vacancies created by the neutron irradiation. Since the rates in Fig. 1 are essentially the same, the vacancies created by electrons appear to be less efficient. This reduced efficiency of ordering could be explained by the assumption that the majority of vacancy-interstitial pairs created by electrons are reasonably close, i.e. the displaced atom does not have enough kinetic energy to travel very far from the vacancy. When they recombine, they annihilate one another. Only a vacancy which escapes from its interstitial can migrate sufficiently, before chance recombination, to cause an appreciable interchange of atoms. Previous measurements on copper irradiated by electrons at low temperatures have shown that the theoretical value for the production of vacancies and interstitials is about twice as large as the experimental results.[9] Although some of the theory is still in question, part of this discrepancy is probably due to the production of adjacent vacancy-interstitial pairs which recombine even at low temperatures. Their motion at low temperatures is possibly due to a lowering of the energy barrier by their associated strain fields. The larger discrepancy reported here suggests that some of the pairs which are far enough apart to be "frozen in" at low temperatures, and thereby detected, are still close enough to recombine rather than escape. The factor fourteen arrived at in these experiments is not strictly interpretable on this simple basis because the enhanced vacancy concentration was assumed to be proportional to flux, which is not correct if the concentration of defects becomes large enough for the annihilation of vacancies by interstitials to be important.

The necessity for the defects to escape recombination in order to contribute to diffusion is demonstrated by another experiment. It was mentioned that when a brass wire was irradiated for 10 days the liquid-nitrogen resistivity decreased 1·7 per cent. In order to observe the effect of concentration of defects on mobility, another brass wire was also irradiated with neutrons for 10 days at liquid-nitrogen temperature and then warmed to room temperature to allow all the defects to move at the same time. The liquid-nitrogen resistivity of this specimen decreased only 0·15 per cent. The difference is due to a reduction of the mean free path of the vacancies arising from the increased concentration of interstitials and hence an increased probability of recombination.

The results of reference 3 for diffusion in α-brass are $H = 39{\cdot}5$ kcal/mole and $\log \tau_0 = -15{\cdot}2$. Using these data in equation (1), the mean atomic jump time for thermal motion at 50°C is 2×10^{11} sec. The data of Fig. 1 obey essentially a first-order rate equation, so that the mean time for an atomic jump can be obtained by taking the time for the resistivity to decrease to $1/e$ of its total decrease. This time is about 6000 sec. This represents a gain in the ordering rate of 3×10^7 by the introduction of excess vacancies by irradiation.

It should be understood that the atoms need to move only a few atomic distances to change the short-range order and that this magnitude of gain is not necessarily applicable to all diffusion-controlled processes. It is clear, however, that changes in order in alloys at temperatures as low as the "freezing in" point of vacancies can now be produced in small samples by electron irradiation.

Acknowledgements—The writer expresses his appreciation to G. J. DIENES and G. H. VINEYARD for helpful discussions, and R. H. SCHULER for his assistance in the electron irradiations.

REFERENCES

1. LI C. Y. and NOWICK A. S. *Phys. Rev.* **103**, 294 (1956).

2. SEITZ F. and KOEHLER J. S. *Solid State Physics* Vol. 2. Academic Press, New York (1956). The reader is referred to this review article as a source of primary references and a review of current knowledge of the field of radiation effects.
3. DAMASK A. C. *J. Appl. Phys.* **27**, 610 (1956).
4. ROSENBLATT D. B., SMOLUCHOWSKI R. and DIENES G. J. *J. Appl. Phys.* **26**, 1044 (1955).
5. ADAM J., GREEN A. and DUGDALE R. A. *Phil. Mag.* **43**, 1216 (1952).
6. DIXON C. E., MEECHAN C. J. and BRINKMAN J. A. *Phil. Mag.* **44**, 449 (1953).
7. DUGDALE R. A. and GREEN A. *Phil. Mag.* **45**, 163 (1954).
8. DIENES G. J. and VINEYARD G. H. To be published.
9. CORBETT J. W., DENNEY J. M., FISKE M. D. and WALKER R. M. *Phys. Rev.* **104**, 851 (1956).

OL.
4
958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 182–189.

A THEORETICAL CALCULATION OF THE RELAXATION OF ATOMS SURROUNDING A VACANCY IN THE BODY-CENTERED CUBIC LATTICE

L. A. GIRIFALCO and J. R. STREETMAN

Lewis Flight Propulsion Laboratory, National Advisory Committee for Aeronautics, Cleveland, Ohio

(*Received* 24 *June* 1957; *revised* 23 *July* 1957)

VOL.
4
1958

Abstract—The extent of relaxation of atoms surrounding a vacancy in body-centered crystals has been computed, assuming an inverse-power pairwise potential acting between atoms in the crystal. Detailed calculations for a 5–7 law were made, using an IBM 650-type electronic computer, which showed that the extent of relaxation was 5·05 per cent for the nearest neighbors and 1·12 per cent for the second-nearest neighbors. Approximate calculations for other potential laws showed that the relaxation is of the order of 3–6 per cent for potential laws ranging from a 4–6 to a 7–12 type.

The energy of vacancy formation in a body-centered lattice whose atoms interact according to the 5–7 potential was found to be 80·1 per cent of the energy of sublimation.

It is concluded from this study that the concept of a *relaxion* as a highly relaxed disordered structure is probably not correct in body-centered crystals whose atoms interact according to an inverse-power potential.

INTRODUCTION

ACCORDING to current theory, diffusion in substitutional alloys and self-diffusion in pure metals proceeds by a vacancy mechanism in which atoms move by jumping into vacant lattice sites that are always in thermodynamic equilibrium with the lattice. The details of the elementary jump mechanism depend on the configuration of the atoms surrounding the vacancy. Thus, if the presence of a vacancy does not alter the positions of adjacent atoms, so that all atoms are at their normal lattice sites, the jump mechanism is quite simple. One of the adjacent atoms moves into the vacancy by squeezing past its neighbors and leaves a vacancy behind it. If, however, the atoms around a vacancy are relaxed inward to a considerable extent, the picture is more complex.

On the basis of diffusion measurements in sodium, it has been suggested[1,2] that the atoms surrounding a vacant site actually exhibit extensive relaxation and that the resulting localized structure, which is called a ***relaxion***, has a density and energy content very similar to that of the liquid phase. It was suggested that the relaxion concept is also valid in other body-centered crystals. The elementary diffusion process is then considered to take place by a kind of local "melting and freezing" in which an atom on the edge of a relaxion moves into it (melting) as another atom moves out of it (freezing). Within a relaxion, atom movements would take place by the co-operative motion of a number of atoms as they roll and squeeze by each other with a rate characteristic of liquid diffusion.

In view of these experiments and their interpretation, it was considered worth while to calculate the extent of vacancy relaxation in order to determine the extent to which theory supported the relaxion concept.

Such calculations have been performed for the face-centered cubic structure by KANZAKI[3] for argon, using a 6–10 potential and by HALL[4] for a Lennard-Jones crystal using a 6–12 potential, but no previous calculations exist for body-centered systems.

The method used in this paper enables calculations to be made for a large number of body-centered systems, but is not applicable to the alkali metals for reasons to be noted in the next section. The theoretical computations of this investigation

can therefore give information concerning relaxion structure for many body-centered cubic crystals, though no definite conclusions can be drawn about its structure in sodium and the other alkali metals.

CHOICE OF A MODEL

The model chosen in this work consists of a body-centered cubic lattice, perfect except for the presence of a single vacancy and whose atoms are all at rest. It is assumed that the entire energy of the crystal can be expressed as a sum of spherically symmetrical interaction potentials between pairs of atoms. The stable relaxion configuration is then obtained by minimizing the energy of interaction between each atom and the rest of the lattice as a function of atom position.

A two-term inverse-power potential function was chosen to represent the energy of interaction of two atoms. Thus, if $\mathscr{E}_{ij}$ is the energy of interaction between the ith and jth atoms that are a distance r_{ij} apart, then

$$\mathscr{E}_{ij} = A\left[\frac{n}{m} {r_0}^{m-n} \frac{1}{{r_{ij}}^m} - \frac{1}{{r_{ij}}^n}\right] \tag{1}$$

where A, n, and m are constants and r_0 is the equilibrium distance of closest approach of two isolated atoms. FÜRTH[5] has correlated the properties of many metals with a function of the form of equation (1) and has computed values of m and n from empirical data.

The first term represents the repulsive potential and the second term the attractive potential between two atoms, so that $m > n$. Pairwise summation of equation (1) over an infinite perfect lattice yields the energy of sublimation, which is finite, so that the sum must converge. This requires that $n > 3$. We therefore have the condition that $m > n > 3$. Unfortunately, this condition is not fulfilled for the alkali metals, and therefore equation (1) is not valid for these materials. The detailed calculations of vacancy relaxation were carried out using equation (1) with $m = 7$ and $n = 5$, which has been used by FÜRTH[5] to describe the pairwise interaction in molybdenum, chromium, and tungsten, so that the results should be valid at least for these metals. Approximate calculations using other values of m and n were also performed.

The following limitations of this method should be kept in mind: equation (1) is a completely empirical function and its use assumes that the cohesive energy can be expressed as a sum of pairwise interactions. This is only approximately true in metals. Furthermore, the constants in equation (1) are obtained from bulk-crystal data and are therefore characteristic of a perfect crystal. It is not known how the electronic redistribution resulting from the removal of an atom to form a vacancy might change the values of m and n.

Equation (1), however, has the advantage that it can be used to compute the extent of vacancy relaxation and the energy of vacancy formation. It turns out that the degree of relaxation is not very sensitive to the values of m and n, so that some confidence can be placed in the results.

LATTICE SUMMATIONS IN THE BODY-CENTERED-CUBIC CRYSTAL

In performing lattice summations it is convenient to express distances in units of a lattice parameter. In terms of the dimensionless distance R_{ij}, which is defined by

$$R_{ij} = \frac{r_{ij}}{a} \tag{2}$$

where a is the half-cell distance (i.e. $1/\sqrt{3}$ times the distance between two nearest neighbors) equation (1) becomes

$$\mathscr{E}_{ij} = A\left[\frac{n}{m} \cdot \frac{{r_0}^{m-n}}{a^m} \cdot \frac{1}{{R_{ij}}^m} - \frac{1}{a^n} \cdot \frac{1}{{R_{ij}}^n}\right]. \tag{3}$$

The energy of sublimation per mole is

$$U = -\frac{AN}{2}\left[\frac{n}{m} \cdot \frac{{r_0}^{m-n}}{a^m} S_m - \frac{1}{a^n} \cdot S_n\right] \tag{4}$$

where

$$S_m = \sum_i \frac{1}{{R_{ij}}^m} \tag{5}$$

$$S_n = \sum_i \frac{1}{{R_{ij}}^m} \tag{6}$$

and N is Avagadro's number. The sums (5) and (6) are carried out for interaction between one atom and the rest of the perfect lattice. They are independent of the value of a, so that equation (4)

gives the energy of the crystal as a function of the lattice spacing. Differentiating equation (4) with respect to a and setting the derivative equal to zero at $a = a_0$ gives a relation between r_0 and the equilibrium lattice spacing:

$$\left(\frac{r_0}{a_0}\right)^{m-n} = \frac{S_n}{S_m}. \tag{7}$$

The energy of any atom in the lattice is obtained by a summation of equation (3). Thus if the energy of the jth atom in a body-centered lattice containing a single vacancy is denoted by $\mathscr{E}_j$, then

$$\mathscr{E}_j = \frac{A}{{a_0}^n}\left[\frac{n}{m}\cdot\frac{S_n}{S_m}\sum_i^{Vj}\frac{1}{{R_{ij}}^m} - \sum_i^{Vj}\frac{1}{{R_{ij}}^n}\right] \tag{8}$$

or, letting

$$S_n{}^{Vj} = \sum_i^{Vj}\frac{1}{{R_{ij}}^n}$$

$$\mathscr{E}_j = \frac{A}{{a_0}^n}\left[\frac{n}{m}\cdot\frac{S_n}{S_m}\cdot S_m{}^{Vj} - S_n{}^{Vj}\right]. \tag{9}$$

The superscript Vj indicates that the summation is performed over a lattice containing a vacancy, taking the jth atom as the origin of co-ordinates. The sums $S_m{}^{Vj}$ and $S_n{}^{Vj}$ depend on the positions of the atoms around the vacancy, so that the stable configuration of the relaxion can be determined by computing $\mathscr{E}_j$ for various atom positions and choosing the configuration at which the energy is a minimum. It is only necessary, of course, to minimize the factor in the brackets of equation (9), so that the relaxion configuration depends only on the choice of m and n.

The configuration of minimum energy can be obtained by a method of successive approximations as follows. Choose a co-ordinate system with origin at the center of the vacancy and axes parallel to the cubic axes of the crystal. Then, in the unrelaxed configuration, the eight nearest neighbors to the vacancy have co-ordinates (1, 1, 1), (1, 1, −1), etc. (distances being measured in units of the half-cell distance) and the six next-nearest neighbors have co-ordinates (0, 0, 2), (0, 2, 0), etc. In the equilibrium configuration, the nearest neighbors are relaxed inwards. Such inward relaxation brings them closer to the second-nearest neighbors, which are therefore repelled and move outward. Since every nearest neighbor has an environment similar to every other nearest neighbor, they all move equal distances along the body diagonal. Similarly, the second-nearest neighbors all move equal distances outward along the X, Y, and Z axes. Therefore, in the relaxed positions, the co-ordinates of the nearest neighbors are $(1-\alpha, 1-\alpha, 1-\alpha)$, $(1-\alpha, 1-\alpha, -1+\alpha)$, etc. and the co-ordinates of the next-nearest neighbors are $(0, 0, 2+\beta)$, $(0, 2+\beta, 0)$, etc.

An estimate of the extent of inward relaxation of the nearest neighbors can be obtained by holding all other atoms in their normal lattice positions and computing the energy of interaction between one of the nearest neighbors and the rest of the lattice as a function of α. The position of minimum energy determines the first approximation to α. Leaving the nearest neighbors in this position, an estimate of β can be obtained by computing the energy of

Table 1(a). *The sums* $S_n{}^{V1}$ *for the configuration in which all atoms except the nearest neighbors to the vacancy are in normal lattice positions*

$1/\sqrt{3}$ (Distance from origin) $= 1-\alpha$	$S_4{}^{V1}$	$S_5{}^{V1}$	$S_6{}^{V1}$	$S_7{}^{V1}$	$S_8{}^{V1}$	$S_9{}^{V1}$	$S_{10}{}^{V1}$
1·05	2·240104	0·886423	0·424729	0·221615	0·120521	0·067627	0·037779
1·0	2·233212	0·877574	0·416517	0·214943	0·115483	0·063394	0·035241
0·95	2·244549	0·882795	0·418697	0·215702	0·115631	0·063317	0·035106
0·9	2·277004	0·904246	0·432743	0·224837	0·121537	0·067110	0·037528
0·8	2·429000	1·016834	0·513918	0·282184	0·161388	0·094419	0·056018
0·7	2·781034	1·305919	0·743949	0·461322	0·298710	0·198416	0·134023
0·6	3·574815	2·044488	1·411754	1·054448	0·819165	0·651100	0·525104

interaction between one of the next-nearest neighbors and the rest of the crystal (including the relaxed nearest neighbors) as a function of β. The position of minimum energy determines the first approximation to β. With this value of β, a second approximation to α can be computed by recalculating the energy of the nearest neighbors as a function of position. The process can be continued until successive approximations differ by an arbitrary small amount. If necessary, the relaxation of third-nearest neighbors can be calculated in the same way.

Values of the sums $S_n{}^{Vj}$ required for these calculations were obtained by means of a type-650 IBM electronic computer. Summations were performed on a finite lattice consisting of a cube $32a_0$ on a side and containing a vacancy in its center. The results are given in Tables 1(a) through 1(e) as a function of atomic position. The sums S_n have been tabulated by JONES and INGHAM[6] and MISRA[7] for various values of n for the simple, face-centered, and body-centered cubic lattices. These sums were also obtained for the body-centered and face-centered structures in the course of the present work, using the electronic computer, and are tabulated in Table 2. The summations were

Table 1(b). *The sums $S_5{}^{V2}$ and $S_7{}^{V2}$ for the configuration in which the first-nearest neighbors are a distance $\sqrt{3}(0{\cdot}953)$ from the origin and all other atoms except the next-nearest neighbors are in normal lattice positions*

Distance from origin $= 2+\beta$	$S_5{}^{V2}$	$S_7{}^{V2}$
2·1	0·937604	0·241548
2·05	0·932386	0·238456
2·025	0·931371	0·237907
2·0	0·931356	0·237992
1·95	0·934540	0·240165

Table 1(c). *The sums $S_5{}^{V1}$ and $S_7{}^{V1}$ for the configuration in which the next-nearest neighbors are at a distance 2·02 from the origin and all other atoms except the first-nearest neighbors are in normal lattice positions*

$1/\sqrt{3}$ (Distance from origin) $= 1-\alpha$	$S_5{}^{V1}$	$S_7{}^{V1}$
0·98	0·870931	0·211138
0·96	0·873145	0·211490
0·95	0·875165	0·212134
0·94	0·877828	0·213108
0·92	0·885188	0·216100

Table 1(d). *The sums $S_5{}^{V2}$ and $S_7{}^{V2}$ for the configuration in which the nearest neighbors are at a distance $\sqrt{3}(0{\cdot}949)$ from the origin and all other atoms except the next-nearest neighbors are at their normal lattice positions*

Distance from origin $= 2+\beta$	$S_5{}^{V2}$	$S_7{}^{V2}$
2·025	0·932944	0·238617
2·0224	0·932905	0·238602
2·02	0·932874	0·238592
2·015	0·932877	0·238607

Table 1(e). *The sums $S_5{}^{V1}$ and $S_7{}^{V1}$ for the configuration in which the next-nearest neighbors are at a distance 2·0224 from the origin and all other atoms except the first-nearest neighbors are in normal lattice positions*

$1/\sqrt{3}$ (Distance from origin) $= 1-\alpha$	$S_5{}^{V1}$	$S_7{}^{V1}$
0·947	0·875018	0·211982
0·948	0·874755	0·211887
0·949	0·874517	0·211806
0·950	0·874273	0·211720

Table 2. *Lattice sums for cubic crystals*

n	S_n (b.c.c.)	S_n (f.c.c.)
4	2·514842	6·322309
5	0·947292	3·000245
6	0·453815	1·806529
7	0·236374	1·180768
8	0·127841	0·800107
9	0·0705263	0·552091
10	0·0393596	0·384723
12	0·0125023	—
14	0·00403143	—

performed over a cube $60a$ on a side. The correction terms obtained by integrating from the faces of this cube to infinity are negligible except for $n = 4$ and $n = 5$. The integration contribution is of the order of 3 per cent for $n = 4$ and 0·2 per cent for $n = 5$. These corrections are included in the tabulated values. The values in Table 2 differ somewhat from those given in references (4) and (5), especially for small values of n. The sums in Table 2 should be more accurate than those given in the previous references, since no mathematical-approximation techniques were employed in the present work.

THE EXTENT OF RELAXATION

The relative displacement of an atom from its normal position resulting from vacancy formation is defined by $\delta = 1 - (d_r/d_u)$, where d_r is the distance from the center of the vacancy to the relaxed atom and d_u is the distance from the center of the vacancy to the atom in its original unrelaxed position. Positive values of the relaxation parameter δ indicate an inward displacement and negative values indicate outward displacement. δ is easily computed from the values of α and β which were obtained by minimizing the energy of equation (9).

Table 3. The degree of relaxation of atoms around a vacancy for an inverse 5–7 interaction potential

	Nearest neighbor δ	Next-nearest neighbor δ
First approximation	0·047	−0·010
Second approximation	0·051	−0·0112
Third approximation	0·0505	—

Table 3 shows the extent of relaxation of the atoms around a vacancy for an interaction potential in which $n = 5$ and $m = 7$. The differences between successive approximations to δ converged rapidly to zero, so that it was necessary to carry out the computation of the extent of relaxation only to the third approximation for nearest neighbors and to the second approximation for next-nearest neighbors. As the table shows, the relaxation is rather small, amounting to only 5 per cent for the nearest neighbors.

The second approximation to the energy of an atom adjacent to the vacant site as a function of $(1-\alpha)$ is shown in Fig. 1.

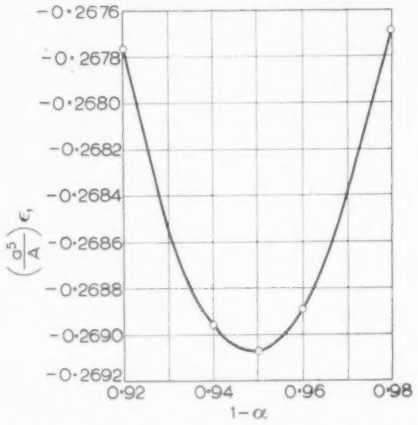


FIG. 1. The energy of an atom adjacent to a vacancy as a function of the extent of relaxation.

It is of interest to investigate the effects of the parameters m and n in the interaction potential on the degree of relaxation. The contribution of the first-nearest neighbors to the forces that determine the equilibrium configuration of the relaxion depends on the distance between them. Thus, if two atoms are further apart than r_0, the pairwise equilibrium distance, they attract each other and inward relaxation is favored. If they are closer together than r_0, they will repel each other and thereby hinder inward relaxation. In the unrelaxed position, the distance between two adjacent atoms that are nearest neighbors to the vacancy is $2a_0$. According to equation (7)

$$\frac{2a_0}{r_0} = 2\left(\frac{S_m}{S_n}\right)^{1/(m-n)}. \tag{10}$$

A ratio of $2a_0/r_0$ greater than unity favors inward relaxation, while a value less than unity hinders it. Values of this ratio have been computed and are listed in Table 4. It is apparent from this table that the degree of relaxation will be less in solids with small values of m and n than in solids with high values of m and n. For $n = 5$ and $m = 7$, the attractive and repulsive forces between two adjacent atoms that are nearest to the vacant site are almost exactly balanced when the atoms are at their normal lattice positions, and become repulsive as

Table 4. Distance between adjacent nearest neighbors for various interaction potentials

n	m	$2a_0/r_0$
4	6	0·850
4	7	0·907
4	8	0·947
4	9	0·979
5	7	0·999
5	8	1·026
5	10	1·059
6	8	1·062
6	9	1·075
6	10	1·085
6	12	1·099
7	10	1·100
7	12	1·111

inward relaxation occurs. This accounts for the fact that the calculated relaxation effect is small.

Even for values of m and n most favorable to inward relaxation, the ratio $2a_0/r_0$ is greater than unity by only about 10 per cent. This means that in any body-centered lattice whose atoms interact according to an inverse-power law, the relaxation of the atoms around a vacancy is small.

A calculation of the first approximation to the relaxation parameter of the nearest neighbors to the vacancy, for various values of m and n, has been performed using the technique outlined above. The results, shown in Table 5, are in agreement with the qualitative conclusions derived from the values of $2a_0/r_0$ in that δ varies only from about 0·03 to 0·06 for all values of m and n, and δ is smaller for lower values of m and n.

Table 5. The first approximation to the relaxation parameter of nearest neighbor for various interaction potentials

n	m	δ
4	6	0·033
4	7	0·037
4	8	0·039
4	9	0·041
5	7	0·047
5	8	0·047
5	9	0·047
5	10	0·048
6	8	0·051
6	9	0·051
6	10	0·052
7	10	0·053

THE ENERGY OF RELAXATION

The degree of relaxation is so small that only the nearest and next-nearest neighbors to the vacancy need be considered in energy calculations. The relaxation energy ΔE_x is therefore given by

$$\Delta E_x = E_R - E_u \tag{11}$$

where E_R is the energy of interaction of the fourteen atoms nearest the vacancy with the rest of the lattice and with each other, when every atom is in its relaxed equilibrium position. E_u is the corresponding energy when every atom is at its original lattice site.

Let $\epsilon_1{}^u$ be the energy of interaction between a nearest-neighbor atom and the rest of the lattice when all atoms are in their normal positions, and let $\epsilon_1{}^R$ be the corresponding energy when all atoms are in their relaxed positions. Let $\epsilon_2{}^u$ and $\epsilon_2{}^R$ be analogous interaction energies for atoms that are second-nearest neighbors to the vacancy. These energies are easily computed from equation (9).

For $n = 5$ and $m = 7$, the energy of interaction between two atoms a distance $R = (x^2+y^2+z^2)^{\frac{1}{2}}$ apart is given by

$$\epsilon(x, y, z) = \frac{A}{a^5}\left[\frac{5}{7}\cdot\frac{S_5}{S_7}\cdot\frac{1}{R^7} - \frac{1}{R^5}\right]. \tag{12}$$

From a consideration of the geometry of the atoms surrounding a vacancy, it is easily seen that in the unrelaxed configuration

$$E_u = 8\epsilon_1{}^u + 6\epsilon_2{}^u - 12\epsilon(0, 0, 2) - 24\epsilon(0, 2, 2) - \\ - 4\epsilon(2, 2, 2) - 24\epsilon(1, 1, 1) - 3\epsilon(0, 0, 4) - 24\epsilon(1, 1, 3) \tag{13}$$

In the relaxed configuration, the first-nearest neighbors have co-ordinates (0·9495, 0·9495, 0·9495), (0·9495, 0·9495, −0·9495), etc. and the second-nearest neighbors have co-ordinates (0, 0, 2·0224, 0), etc. so that the energy of the relaxed configuration is given by

$$E_R = 8\epsilon_1{}^R + 6\epsilon_2{}^R - 12\epsilon(0, 0, 1{\cdot}899) - \\ -12\epsilon(0, 1{\cdot}899, 1{\cdot}899) - \\ -4\epsilon(1{\cdot}899, 1{\cdot}899, 1{\cdot}899) - \\ 24\epsilon(0{\cdot}9495, 0{\cdot}9495, 1{\cdot}0739) - \\ -12\epsilon(0, 2{\cdot}0224, 2{\cdot}0224) - 3\epsilon(0, 0, 4{\cdot}0448) - \\ -24\epsilon(0{\cdot}9495, 0{\cdot}9495, 0{\cdot}9495). \tag{14}$$

Since the sublimation energy per atom is given by equations (4) and (7) as

$$\Delta E_{\text{vap}} = \frac{A}{7a^5} \cdot S_5, \tag{15}$$

the ratio of the relaxation energy to the vaporization energy is

$$\frac{\Delta E_x}{\Delta E_{\text{vap}}} = \frac{7(E_R - E_u)}{[A/a^5]S_5}. \tag{16}$$

$(E_R - E_u)$ contains the factor A/a^5, so that it is not necessary to compute this quantity in order to obtain $\Delta E_x/\Delta E_{\text{vap}}$. With the aid of the electronic computer, calculation of $\Delta E_x/\Delta E_{\text{vap}}$ is quite straightforward and the result is

$$\frac{\Delta E_x}{\Delta E_{\text{vap}}} = -0{\cdot}19916. \tag{17}$$

The total energy of vacancy formation is

$$\Delta E^V = \Delta E_{\text{vap}} + \Delta E_x = \Delta E_{\text{vap}} \left[1 + \frac{\Delta E_x}{\Delta E_{\text{vap}}}\right] \tag{18}$$

so that

$$\Delta E^V = 0{\cdot}8008 \Delta E_{\text{vap}}.$$

SUMMARY AND CONCLUSIONS

Detailed calculations of the extent of relaxation of atoms around a vacancy and of the energy of vacancy formation in a body-centered cubic lattice have been performed, using a type-650 IBM electronic computer. It was assumed that the energy of the crystal could be expressed as a sum of two-term inverse-power pairwise potentials.

The most extensive calculations were performed with an inverse 5–7 interaction potential, which is at least approximately valid for molybdenum, chromium and tungsten. For this type of interaction, the nearest neighbors to the vacancy moved inward by 5·05 per cent of their distance from the vacancy, and the second-nearest neighbors moved outward 1·12 per cent. The relaxation energy reduced the energy of vacancy formation by 19·9 per cent of the energy of vaporization, the total energy of formation being 80·1 per cent of the vaporization energy.

Approximate calculations of the extent of relaxation for pairwise interaction potentials with different values of the exponents showed that the relaxation is not very sensitive to the force law used. The relaxation of first neighbors ranged from 3.3 per cent for a 4–6 law to 5·3 per cent for a 7–10 law.

The concept of a relaxion as a highly relaxed, disordered structure was advanced by NACHTRIEB on the basis of experiments performed at elevated temperatures. Since the calculations presented in this paper refer to absolute zero, they are not quantitatively comparable to high-temperature data. Furthermore, the metal investigated in greatest detail by NACHTRIEB was sodium, and it is known that an inverse-power pairwise potential cannot be used for this system. In view of these facts, no definite conclusions regarding the structure of the relaxion in alkali metals can be given on the basis of the present calculations. However, for body-centered cubic systems that can be described by an inverse-power force law, the theoretical calculations predict that the relaxation around a vacancy is small and that it is unlikely that self-diffusion occurs by the relaxion mechanism.

The calculations of KANZAKI[3] and HALL[4] indicate that the relaxation effect is even smaller for face-centered lattices than for body-centered lattices.

KANZAKI's method of calculations consisted of applying an external distribution of forces to the crystal which results in atomic displacements equivalent to the displacements obtained on removing an atom from the lattice. He assumed a model for solid argon in which each atom is in equilibrium under the effect of the repulsive interaction of the twelve nearest neighbors and the attractive interaction of the six second-nearest neighbors. With an inverse 6–10 potential, this model results in an inward displacement of the atoms adjacent to the vacancy of only 0·3 per cent.

HALL used a relaxation method with a 6–12

potential in which the interactions were not restricted to a small number of neighbors, and arrived at a larger value for the relaxation, the nearest neighbors moving in about 0·8 per cent. The discrepancy between KANZAKI's and HALL's results most probably arises from the fact that KANZAKI considered only first and second neighbors as interacting rather than from the difference in the power of the repulsive part of the interaction.

The method used in the present study is simpler and more direct in principle than either the KANZAKI or the HALL method, and is easily applied to the calculation of the relaxation near the core of the vacancy. However, if a high degree of accuracy in calculating the distortion at large distances from the vacancy is required, the iterative summation procedures becomes quite cumbersome and KANZAKI's or HALL's method may be more convenient.

OL.
4
958

REFERENCES

1. NACHTRIEB N. H. and HANDLER G. S. *Acta Met.* **2,** 797 (1954).
2. NACHTRIEB N. H. *WADC Tech. Rep.* 55–68, University of Chicago (1954).
3. KANZAKI H. *J. Phys. Chem. Solids* **2,** 24 (1957).
4. HALL G. L. *J. Phys. Chem. Solids* **3,** 210 (1957).
5. FÜRTH R. *Proc. Roy. Soc.* A **183,** 87 (1944).
6. JONES J. E. and INGHAM A. E. *Proc. Roy. Soc.* A **107,** 636 (1925).
7. MISRA R. D. *Proc. Camb. Phil. Soc.* **36,** 173 (1940).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 190–195.

DEVIATIONS FROM STOICHIOMETRY IN BINARY IONIC CRYSTALS

ROBERT F. BREBRICK

U.S. Naval Ordnance Laboratory, White Oak, Maryland

(*Received* 21 *June* 1957)

Abstract—A grand partition function has been constructed for a pure uni-univalent ionic crystal which can deviate from stoichiometric composition by the inclusion of lattice vacancies. A departure from earlier statistical treatments is made in assuming that the valence electrons distribute themselves over the energy levels of a typical semiconductor electronic band structure. Each vacancy in the cation (anion) sublattice of the crystal has associated with it an acceptor (donor) level in the band structure with a maximum occupancy of one and a degeneracy of two. The interaction between deviations from stoichiometry and the electronic band structure is assumed to be entirely accounted for by the introduction or removal of these donor and acceptor levels. The equilibrium distribution functions obtained are put in the form of mass-action-law expressions. However, the equilibrium constants for those mass-action-law expressions which contain the concentrations of conduction band electrons or valence band holes are not true thermodynamic quantities. These constants depend on the Fermi level and therefore on the deviation from stoichiometry. This dependence disappears when the Fermi level is in the forbidden gap and a few units of kT away from the conduction and valence bands. The chemical potentials of the crystal components are obtained. These agree simultaneously with the expressions used to date only when the Fermi level is between the donor and acceptor levels and a few units of kT away from them.

VOL.
4
1958

INTRODUCTION

THE theoretical framework for understanding deviations from stoichiometry in ionic crystals was first laid by WAGNER and SCHOTTKY.[1,2] They found that the results of their statistical-mechanical treatment could also be obtained by the application of the mass-action law to various quasichemical reactions in which lattice vacancies, interstitial ions, electrons, etc. appear as chemical species. Recently the theory has been extended heuristically and found to give a satisfactory description of the dependence of the electrical and optical properties of CdS crystals on the conditions of preparation.[3] The deviation from stoichiometry in PbS as a function of the temperature and the pressure of the sulfur vapor ambient at equilibrium has also been extensively studied experimentally[4,5] and interpreted by means of the extended theory.[5] The theoretical basis for the mass-action laws, however, is still that supplied by WAGNER and SCHOTTKY. This basis will be reexamined here under the assumption that the valence electrons of the crystal are distributed over a typical semiconductor electronic band structure. Such an assumption has been made in the analyses of donor and acceptor substitutional impurities in semiconductors recently given by REISS[6] and by LONGINI and GREENE.[7] However, the problem of deviations from stoichiometry and the presentation given here are different enough from those analyses to merit a separate discussion.

THE GRAND PARTITION FUNCTION

For simplicity we shall deal with a binary, uni-univalent crystal, CA. It is assumed that vacancies in the cation and anion sublattices are the only types of defect and that there is one donor (acceptor) level with a maximum occupancy of one electron and a degeneracy of two in the electronic energy-band system for each anion (cation) vacancy. Two fundamental assumptions in the statistical-mechanical treatment of this model are that:

(a) The valence electrons and the atomic kernels

can be treated independently. In particular, the effect of deviations from stoichiometry on the electronic energy-band structure is entirely accounted for by the introduction or removal of the proper number and type of localized levels in the forbidden gap. This assumption has been discussed elsewhere.[6,8]

(b) The concentration of lattice vacancies is small enough for any interaction tending towards clustering to be negligible.

The internal variables denoting the state of the crystal are the following concentrations: N_C, the concentration of cations; N_A, the concentration of anions; N_{hC}, the concentration of cation vacancies; N_{hA}, the concentration of anion vacancies; n_{hA}, the concentration of electrons in donor levels associated with anion vacancies; p_{hC}, the concentration of holes in acceptor levels associated with cation vacancies; n_c, the concentration of electrons in the conduction band; and p_v, the concentration of holes in the valence band. These quantities are related by the equation for the conservation of lattice sites

$$\begin{aligned} N_C &= S-N_{hC} \\ N_A &= S-N_{hA} \end{aligned} \tag{1}$$

where S is the concentration of lattice sites in each sublattice, and by the equation for electroneutrality

$$N_C+p_v+p_{hC} = N_A+n_c+n_{hA}. \tag{2}$$

Taking the standard state as that of the atoms at infinite dispersion, we then write the quasi-grand partition function[9] for the crystal as:

$$\begin{aligned} &(\text{G.P.F.}) = \\ &\sum_K \left\{ \binom{S}{N_{hC}} \binom{S}{N_{hA}} \binom{N_{hC}}{p_{hC}} \binom{N_{hA}}{n_{hA}} \binom{g_v}{p_v} \binom{g_c}{n_c} \times \right. \\ &\times 2(p_{hC}+n_{hA}) \exp[(-N_{hC}\zeta_{hC}-N_{hA}\zeta_{hA}+p_{hC}E_{hc}- \\ &-n_{hA}E_{hA}-n_cE_c+p_vE_v)/kT] \times \\ &\times (\lambda_m K_C \exp -U/kT)^{(S-N_{hC})} \times \\ &\left. \times (\lambda_x K_A \exp -U/kT)^{(S-N_{hA})} \right\} \end{aligned} \tag{3}$$

where the external variables of the crystal are: T, the absolute temperature; λ_m, the absolute activity per atom of the metallic constituent which forms the cations in the crystal; and λ_x, that of the non-metallic constituent. The parameters are: the energies specifying the electronic band structure E_c, E_v, E_{hC}, E_{hA}, which are respectively the energy of the bottom edge of the conduction band, that of the top edge of the valence band, that of the acceptor levels, and that of the donor levels; g_v and g_c, the effective density of states near the top of the valence band and near the bottom of the conduction band respectively; ζ_{hC}, an energy associated with the creation of a cation vacancy by moving a cation from a normal site to one on the surface of the crystal; ζ_{hA}, the analogous quantity for the creation of an anion vacancy; K_C and K_A, the contribution of the cations and anions respectively to the vibrational modes of the crystal; K, that part of the partition function which does not depend on the internal variables listed above; and U is one-half of the crystal lattice energy per cation–anion pair relative to the infinitely dispersed atoms. The quantities $\binom{x}{y}$ are the combinatorial coefficients $x!/y!(x-y)!$.

The sum in equation (3) is over all values of the internal variables consistent with equation (2). At equilibrium the state of the crystal is given by those values of the internal variables that make the grand partition function a maximum. A common procedure is followed in determining these values. The logarithm of the partition function is approximated by that of its largest term. This in turn is determined by maximizing the logarithm of the general term of the sum in equation (3) with respect to each of the internal variables after applying Stirling's approximation for the factorials. The restriction on these variables expressed in equation (2) introduces a Lagrangian multiplier, the Fermi level divided by kT, $(E_{f/kT})$, into the resulting equations. This may be eliminated between the equations themselves and equation (2). The equilibrium values of the internal variables are then obtained as a function of the parameters of the crystal and the external variables λ_m, λ_x, and T. If the crystal is considered to be in equilibrium with its vapor, there is a relation among these variables.[9]

$$Pv = kT \ln (\text{G.P.F.}) \tag{4}$$

where P is the pressure of the vapor phase and v is the crystal volume containing S lattice sites (the unit volume here). The equilibrium crystal-vapor

system therefore has two degrees of freedom, in agreement with the Gibbs phase rule.

ELECTRONIC EQUILIBRIUM

With many materials the electrons distribute themselves over the available energy levels and attain equilibrium rapidly. The equilibrium concentration of lattice defects, however, involves the movement of ions and is not attained so rapidly. As a result the state of a crystal often corresponds to frozen-in, high-temperature distribution of lattice defects and the prevailing lower temperature. The equations describing such a situation are obtained by considering N_C and N_A, and therefore N_{hC} and N_{hA} (equation (1)), fixed in the grand partition function and maximizing it with respect to each of the remaining internal variables. In this way the following equations are obtained:

$$g_c/n_c = 1+\exp[(E_c-E_f)/kT], \tag{5a}$$

$$N_{hA}/n_{hA} = 1+\tfrac{1}{2}\exp[(E_{hA}-E_f)/kT], \tag{5b}$$

$$N_{hC}/p_{hC} = 1+\tfrac{1}{2}\exp[(E_f-E_{hC})/kT], \tag{5c}$$

$$g_v/p_v = 1+\exp[(E_f-E_v)/kT]. \tag{5d}$$

These are the usual electron and hole distribution functions.

LATTICE AND ELECTRONIC EQUILIBRIUM

Maximizing the grand partition function with respect to all of the internal variables, subject to the electrical neutrality condition, one obtains unchanged the electron and hole distribution functions given in equations (5a–d). In addition, the distribution functions for the lattice vacancies are obtained:

$$\frac{N_{hC}}{S-N_{hC}} = \frac{N_{hC}}{N_{hC}-p_{hC}}(\lambda_m K_C)^{-1}\times \times\exp[(E_f-\zeta_{hC}+U)/kT], \tag{5e}$$

$$\frac{N_{hA}}{S-N_{hA}} = \frac{N_{hA}}{N_{hA}-n_{hA}}(\lambda_x K_A)^{-1}\times \times\exp[(-E_f-\zeta_{hA}+U)/kT]. \tag{5f}$$

Equations (2) and (5) specify the equilibrium state of the crystal as a function of the external variables T, λ_m, and λ_x.

MASS-ACTION LAWS

The extent to which mass-action laws are valid for the particular crystal model considered here can now be discussed. For this purpose the following changes in symbols are made:

$$N_{hA} = V_A+V_A^-,\ n_{hA} = V_A^- \tag{6a}$$

$$N_{hC} = V_C+V_C^+,\ p_{hC} = V_C^+. \tag{6b}$$

Using equations (6), the equilibrium distributions given by equations (5) can be rewritten as:

$$n_c = g_c\{1+\exp[(E_c-E_f)/kT]\}^{-1} \tag{7a}$$

$$V_A^-/V_A = 2\exp[(E_f-E_{hA})/kT] \tag{7b}$$

$$V_C^+/V_C = 2\exp[(E_{hC}-E_f)/kT] \tag{7c}$$

$$p_v = g_v\{1+\exp[(E_f-E_v)/kT]\}^{-1} \tag{7d}$$

$$V_C = N_C(\lambda_m K_C)^{-1}\exp[(E_f-\zeta_{hC}+U)/kT] \tag{7e}$$

$$V_A = N_A(\lambda_x K_A)^{-1}\exp[(-E_f-\zeta_{hA}+U)/kT]. \tag{7f}$$

Combinations of equations (7) can be made which under certain conditions are of the form of mass-action laws. Thus equations (8a) through (8e) are respectively obtained by multiplying equation (7c) by (7e), (7b) by (7f), (7e) by (7f), (7a) by (7d), and (7b) by (7d):

$$V_C^+\lambda_m = 2N_C K_C^{-1}\exp[(E_{hC}-\zeta_{hC}+U)/kT]; \tag{8a}$$

$$V_A^-\lambda_x = 2N_A K_A^{-1}\exp[(-E_{hA}-\zeta_{hA}+U)/kT], \tag{8b}$$

$$V_C V_A = N_C N_A(\lambda_m\lambda_x K_C K_A)^{-1}\times \times\exp[(-\zeta_{hC}-\zeta_{hA}+2U)/kT] \tag{8c}$$

$$n_c p_v = g_c g_v\left[1+\exp\left(\frac{E_c-E_f}{kT}\right)\right]^{-1}\times \times\left[1+\exp\left(\frac{E_f-E_v}{kT}\right)\right]^{-1} \tag{8d}$$

$$\frac{V_A^- p_v}{V_A} = 2g_v\left[1+\exp\left(\frac{E_f-E_v}{kT}\right)\right]^{-1}\times \times\exp\left(\frac{E_f-E_{hA}}{kT}\right). \tag{8e}$$

These equations are to be compared with the mass-action law expressions below:

VOL. 4 1958

$$V_C{}^{+}\lambda_x{}^{-1} = k_1(T), \tag{9a}$$

$$V_A{}^{-}\lambda_m{}^{-1} = k_2(T), \tag{9b}$$

$$V_C V_A = k_3(T), \tag{9c}$$

$$n_c p_v = k_4(T), \tag{9d}$$

$$V_A{}^{-} p_v / V_A = k_5(T), \tag{9e}$$

where the k's are functions of temperature alone. Equations (9a) and (9b) represent the interaction of the crystal with a vapor phase of its components, leading to deviations from stoichiometry. Equations (9c) and (9d) represent the intrinsic formation of Schottky defects and that of conduction-band electrons and valence-band holes respectively. Equation (9e) represents the combination of a hole in the valence band with an electron in a donor level.

The exact equations (8a) through (8e) reduce to the mass-action laws (9a) through (9e) if:

(1) The deviations from stoichiometry are small enough for the relative change in N_C and N_A to be neglected.

(2) The product of the absolute activities is a function of temperature only.

(3) The Fermi level is far enough away from the bands for classical statistics to be applicable to the electrons in the conduction band and the holes in the valence band.

The first condition is necessary for the validity of the statistical treatment itself and is assumed to hold.

Using the definition of the absolute activity

$$\lambda = \exp(\mu/kT) \tag{10}$$

where μ is the chemical potential, the second condition is seen to be equivalent to

$$\mu_m + \mu_x = k(T). \tag{11}$$

Assuming the vapor phase in equilibrium with the crystal is ideal, the equation

$$\mu_i = kT \ln p_i + \mu_i{}^0(T) \tag{12}$$

holds and condition (2) is equivalent to

$$p_m p_x = k^1(T). \tag{13}$$

Equations similar to (13) are commonly assumed to hold in the description of deviations from stoichiometry by the use of the mass-action law, and reduce the number of independent external variables to two. It can be shown quite generally by a purely thermodynamic argument, using the Gibbs–Duhem relation, that the molar Gibbs free energy is given by equation (11), provided that the deviations from stoichiometry are small; $N_{hC}/S \ll 1$ and $N_{hA}/S \ll 1$. Thus the free energy is given by

$$\begin{aligned} F &= 2N_C\mu_m/(N_C+N_A)+2N_A\mu_x/(N_C+N_A) \\ &= \mu_m+\mu_x+(N_{hA}-N_{hC})(\mu_m-\mu_x)/S \\ &\cong \mu_m+\mu_x. \end{aligned} \tag{14}$$

That the sum of the chemical potentials is a function of temperature can be shown only by the Gibbs–Duhem relation

$$d\mu_m = -(N_A/_{\perp}\ _C)d\mu_x \text{ (const. } T, P). \tag{15}$$

On substituting for N_C and N_A from equation (1), this equation becomes

$$d\mu_m = -\left(\frac{1-N_{hA}/S}{1-N_{hC}/S}\right)d\mu_x \cong -d\mu_x. \tag{16}$$

It only remains to show that the statistical treatment here is consistent with this result. This can be done with the aid of equation (4) and the equilibrium distribution functions given by equations (7), but will not be carried through here. The significance of equations (11) and (13) for binary crystals showing small deviations from stoichiometry has been appreciated earlier by WAGNER[1,10] and others.[3,5]

The fulfilment of the first two conditions is sufficient for the reduction of equations (8a), (8b), and (8c) to mass-action-law expressions. The third condition must also hold for the reduction of equations (8d) and (8e), however. The circumstances under which this is so depend not only on the deviation from stoichiometry but also on the temperature, the width of the energy gap in the band structure, and the density and position of the "impurity" levels.

In the statistical treatment given here, the six internal variables and the Fermi level are determined by the six equations (7a) through (7f) and the electrical-neutrality condition (equation (2)) as a function of the external variables. Mass-action-law expressions, which are equations giving the ratio of simple products of the internal variables as a function of temperature, can be obtained from

certain combinations of equations (7) when classical statistics are applicable to the conduction-band electrons and valence-band holes. When this is so, the Fermi level appears as a power of an exponential factor in equation (7), and it is only necessary to choose those combinations in which the Fermi level cancels out. In three of the six equations the Fermi level appears with a plus sign and in the other three with a negative sign. Therefore a total of nine equations of the form $xy = k(T)$ and six of the form $x/y = k(T)$ can be obtained. Only five of these are independent, however, and equations (8a) through (8e) comprise such an independent set. Thus these five equations and the electrical-neutrality condition allow one to solve for the internal variables as a function of external variables. In this procedure the Fermi level has been purposely eliminated and cannot be obtained.

It is interesting to note that any quasi-chemical reaction that can be written involving the internal variables of the crystal leads to a mass-action-law expression that is derivable from equation (7) in the limit of classical statistics. This is due to the way in which the Fermi level appears in these equations and the restriction that an acceptable chemical reaction be electrically balanced.

Finally, it is to be expected that the general results for the particular crystal model considered here will also follow when other types of point defects and more complicated electronic energy-band structures are considered.

CHEMICAL POTENTIALS OF THE CRYSTAL COMPONENTS

Using the general relation between the absolute activity and the chemical potential given in equation (10), the chemical potentials of the two crystal components can be directly obtained from equations (5e) and (5f). These are:

$$\mu_m = kT\ln(N_C/N_{hC}) - kT\ln(1-p_{hC}/N_{hC}) + \\ + E_f - \zeta_{hC} + U - kT\ln K_C \qquad (17a)$$

$$\mu_x = kT\ln(N_A/N_{hA}) - kT\ln(1-n_{hA}/N_{hA}) - \\ - E_f - \zeta_{hA} + U - kT\ln K_A. \qquad (17b)$$

Only the first three terms of the right-hand members of these equations depend on the internal variables of the crystal. Of these the last two arise from the electronic-band scheme. They are not present in the statistical treatments of deviations from stoichiometry ordinarily given, in which the valence electrons are assumed to occupy atomic or ionic levels. The chemical potentials depend on the internal variables through terms exactly analogous to those appearing in the expressions for the chemical potentials of donor and acceptor substitutional impurities.[7] A difference arises in the signs before these terms because the donor and acceptor levels are associated with lattice vacancies in the case of deviations from stoichiometry and with the donor and acceptor atoms themselves in the case of substitutional impurities.

With the aid of equations (6), equations (17) may be rewritten as:

$$\mu_m = kT\ln(N_C/V_C) + E_f - \zeta_{hC} + U - kT\ln K_C \qquad (18a)$$

$$\mu_x = kT\ln(N_A/V_A) - E_f - \zeta_{hA} + U - kT\ln K_A. \qquad (18b)$$

From these equations one sees that at a given temperature the chemical potentials can vary over an energy range at least as large as the separation between the valence and conduction bands, provided that the range of stability of the crystal is wide enough. On the other hand the molar Gibbs free energy is given by the sum of the chemical potentials, and this sum is a function only of temperature to a good approximation, as seen from equations (14) and (16). From equations (14) and (18), the free energy is given by:

$$F = kT\ln(N_C N_A/V_C V_A) - \zeta_{hC} - \zeta_{hA} + \\ + 2U - kT\ln K_C K_A. \qquad (19)$$

Finally, from equations (18), the chemical potential of the metallic (non-metallic) component can be considered as the sum of that of an electron (hole) plus that of a cation (anion). The chemical potentials of the cation and anion depend only on the concentrations of ionized vacancies, however. When the Fermi level is between the donor and acceptor levels and a few units of kT away from them, almost all of the vacancies are ionized and their concentrations may be approximated by the total concentrations of vacancies.

SUMMARY

A statistical-mechanical treatment of a binary

crystal showing deviations from stoichiometry has been given under the assumption that the valence electrons of the crystal are distributed over a typical semiconductor band structure. The specification of the internal state of the crystal by mass-action-law expressions is found to be valid provided that the deviations from stoichiometry are small. For those expressions containing the concentrations of conduction-band electrons or valence-band holes, it is further required that classical statistics be applicable to these quantities.

It is shown that the chemical potentials of the uni-univalent crystal components (equation (18)) can be considered to be the sum of that for an electron, E_f, and that of an ion. It is to be noted that it is the concentration of ionized vacancies rather than the total concentration that appears in the latter.

REFERENCES

1. WAGNER C. and SCHOTTKY W. *Z. Phys. Chem.* B **11,** 163 (1931).
   WAGNER C. *Ibid.* B **22,** 181 (1933).
   WAGNER C. *Z. Electrochem.* **39,** 543 (1935).
2. SCHOTTKY W. *Z. Phys. Chem.* B **29,** 335 (1935).
3(a). KROGER F. A., VINK H. J. and VAN DEN BOOMGAARD J. *Z. Phys. Chem.* **203,** 1 (1954).
   (b). For a general treatment of the mass-action-law approach see KROGER F. A. and VINK H. J. *Solid State Physics* (Edited by F. SEITZ and D. TURNBULL) Vol. 3, pp. 310, 438. Academic Press, New York (1956).
4. BREBRICK R. F. and SCANLON W. W. *Phys. Rev.* **96,** 598 (1954).
5. BLOEM J. and KROGER F. A. *Z. Phys. Chem.* (*Frankfurt*) **7** (1956).
6. REISS H. *J. Chem. Phys.* **21,** 1209 (1953).
7. LONGINI R. L. and GREENE R. F. *Phys. Rev.* **102,** 992 (1956).
8. JAMES H. M. *Photoconductivity Conference* (Edited by R. BRECKENRIDGE, B. R. RUSSELL and E. E. HAHN) p. 204. John Wiley, New York (1956).
9. RUSHBROOKE G. S. *Introduction to Statistical Mechanics* Chap. 17. University Press, Oxford (1949); FOWLER R. H. and GUGGENHEIM E. A. *Statistical Thermo-dynamics* Chap. 6. University Press, Cambridge (1949).
10. WAGNER C. *J. Chem. Phys.* **21,** 1819 (1953).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 196–200.

SELF-DIFFUSION AND ELECTRICAL CONDUCTIVITY IN SILVER BROMIDE*

ALLAN S. MILLER† and ROBERT J. MAURER

Department of Physics, University of Illinois, Urbana, Illinois

(*Received* 12 *August* 1957)

Abstract—The ionic conductivity and the self-diffusion coefficient of silver have been measured in silver bromide and in silver bromide–cadmium bromide single crystals between 20 and 380°C. There exists a range of temperature for crystals containing cadmium within which the ratio of the conductivity to the self-diffusion coefficient has the value of 1·25 which is predicted for the vacancy mechanism of diffusion by the Einstein relation as modified by the Bardeen–Herring correction. This result is in agreement with the assumption that divalent cadmium ions replace monovalent silver ions in the crystal lattice and introduce an equal number of silver-ion vacancies. These excess vacancies reduce the concentration of interstitial silver ions to a negligible quantity through the operation of the mass-action law. Under these conditions diffusion occurs only by the vacancy mechanism. Conductivity and diffusion data obtained with pure silver bromide crystals are in fair agreement with the more extensive data of FRIAUF.

VOL
4
195

INTRODUCTION

THE measured self-diffusion coefficients of silver in silver chloride[1] and in silver bromide[2] are significantly smaller than predicted by the measured conductivity and the Einstein relation.‡

$$D = \sigma\left(\frac{kT}{Ne^2}\right). \qquad (1)$$

This discrepancy has been attributed by MCCOMBIE and LIDIARD[4] to the existence of Frenkel disorder and the interstitialcy mechanism of migration of interstitial silver ions in silver chloride and bromide.[5,7] An interstitial silver ion that migrates by the interstitialcy mechanism jumps into a neighboring lattice site by displacing the silver ion that has occupied the site into an adjacent interstitial position.

Successive interstitialcy jumps are correlated and equation (1) should be corrected by the BARDEEN–HERRING correlation factor.[8,10] A second and larger correction to equation (1) is necessary because the ionic charge suffers a greater displacement than that of either of the two ions participating in an interstitialcy jump. Fig. 1 illustrates a collinear interstitialcy jump in which the displacement of charge is twice that of either ion.

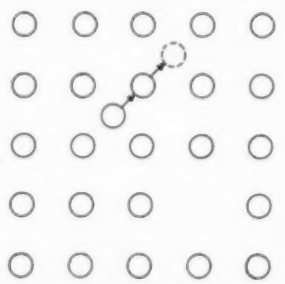


FIG. 1. Schematic representation of Frenkel disorder in silver bromide. A collinear interstitialcy jump is illustrated.

The modification of the Einstein equation (1) by the assumption of collinear interstitialcy jumps yields at heoretical conductivity–diffusion-coefficient ratio that is significantly *larger* than is found

* Based on a Ph.D thesis submitted to the University of Illinois. Partially supported by the Office of Scientific Research of the U.S. Air Force and the National Science Foundation.

† Now at IBM, Poughkeepsie, New York.

‡ The contribution of the negative ion to the conductivity is negligible in these compounds.[1,3] The electronic conductivity is also negligible.

experimentally. The predicted value of this ratio in silver chloride at 350°C is $2{\cdot}34(Ne^2/kT)$. COMPTON and MAURER[1] measured a ratio of $1{\cdot}78(Ne^2/kT)$. This predicted value includes the effect of the contribution to the conductivity and the diffusion of the silver-ion vacancies which are present in the pure crystal with a concentration equal to that of the interstitial silver ions if Frenkel disorder of the lattice is assumed. The relative contributions of the interstitial silver ions and silver-ion vacancies to the conductivity can be calculated from the data of TELTOW[6] and EBERT and TELTOW.[11]

FRIAUF[2] has ascribed the difference between the observed conductivity–diffusion-coefficient ratio and that predicted by the collinear interstitialcy mechanism to neglect of noncollinear jumps. The theoretical calculations of HOVE[12] indicate that a direct jump of a silver ion from an interstitial position to an adjacent interstitial position is relatively improbable.

If a divalent cadmium ion is substituted for silver in the silver bromide lattice, the concentration of silver-ion vacancies is much larger than that of interstitial silver ions at a temperature such that the concentration of Frenkel defects in the pure crystal is small relative to the concentration of cadmium ions. Each divalent cadmium ion introduces a vacancy into the crystal in order to preserve electrical neutrality and the mass-action law demands that the product of the concentration of vacancies and interstitials be a constant at constant temperature. Under this circumstance the migration of the silver ions is determined by the properties of the vacancies. The Einstein relation needs correction only by the Bardeen–Herring correlation factor for the vacancy mechanism of self-diffusion in a face-centered lattice. COMPTON and MAURER[1] reported a conductivity–diffusion-coefficient ratio of approximately $1{\cdot}2(Ne^2/kT)$ in a silver chloride–cadmium chloride crystal. The theoretical value of this ratio for vacancy diffusion is $1{\cdot}25(Ne^2/kT)$. The primary purpose of the present paper is to present similar but more extensive data for silver bromide–cadmium bromide crystals.

EXPERIMENTAL PROCEDURE

Silver bromide was precipitated from a 0·1-N solution of high-purity DuPont silver nitrate with redistilled DuPont reagent grade 0·1-N hydrogen bromide solution. The precipitate was washed, filtered, and vacuum dried. Crystals were also grown from Goldsmith Brothers' reagent-grade silver bromide. Spectroscopic analysis of the crystals indicated that the concentration of iron and copper was 0·001 per cent in the DuPont silver bromide and ten times larger in the Goldsmith silver bromide. The concentration of magnesium was 0·0001 per cent in each sample.

Crystals were grown by the Bridgman technique in Vycor crucibles. The silver bromide was premelted in vacuum, and the Vycor crucible was sealed in a Pyrex capsule under a helium pressure of 300 mm Hg. Silver bromide–cadmium bromide crystals were grown by adding reagent-grade cadmium bromide to the melt of a pure silver bromide crystal. The concentration of cadmium in the crystals was determined by polarographic analysis.

The methods of measurement of conductivity and diffusion were similar to those described by COMPTON and MAURER.[1] Conductivity samples were approximately $8\times8\times4$ mm. Silver paint or Aquadag electrodes were applied to the ground and polished faces of the sample, which was supported between platinum electrodes in a helium-filled oven during measurement. The sample temperature was obtained with a calibrated platinum–platinum 10 per cent rhodium thermocouple.

The diffusion samples were cylinders approximately 1 cm in length and 1 cm in diameter. They were annealed for 30 min at 400°C before use. ^{110}Ag was chemically deposited on a plane face that had been prepared by cutting with a microtome knife. The ^{110}Ag was transformed to silver bromide by immersion in bromine water. The thickness of this layer of silver bromide was approximately 1μ. After diffusion of the ^{110}Ag, the samples were cut into sections approximately 6μ thick with the microtome, and the radioactivity of each section was determined with the equipment previously described.[13]

The diffusion coefficient, D, of the ^{110}Ag was obtained from the plot of the logarithm of the radioactivity per unit thickness of a section as a function of the square of the distance, x^2, of the section from the face of the diffusion sample. Such plots yielded straight lines, as is consistent with the expectation that the concentration $c(x, t)$ of ^{110}Ag in the crystal, after diffusion has occurred for a time t at temperature T, is given by

$$c(x,t) = \frac{c_0}{\sqrt{(\pi Dt)}} \cdot \exp\left(-\frac{x^2}{4Dt}\right). \qquad (1)$$

The measured conductivities and diffusion coefficients are believed to be accurate to ±5 per cent.

RESULTS

The conductivity of the crystals grown from high-purity and reagent-grade silver bromide is shown as a function of temperature in Fig. 2. The data of TELTOW,[6] LUCKEY and WEST,[14] and

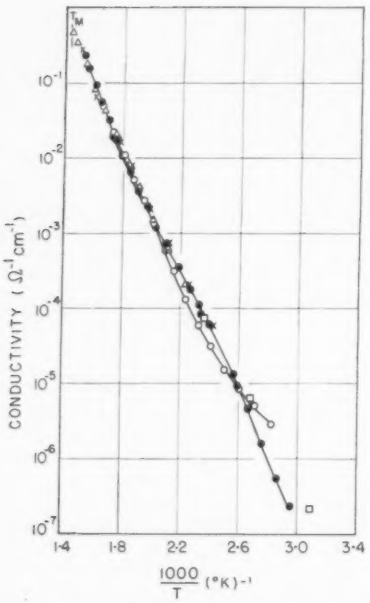


FIG. 2. The conductivity of "pure" silver bromide. ● High purity; ○ reagent grade; △ TELTOW;[6] □ LUCKEY and WEST;[14] × FRIAUF.[2]

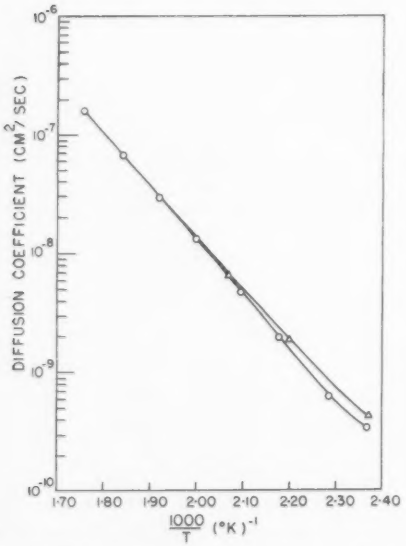


FIG. 3. The self-diffusion coefficient of silver in silver bromide. △ High purity; ○ reagent grade.

Table 1. Electrical conductivity and self-diffusion coefficient of silver bromide at various temperatures.

Temperature (°C)	Conductivity ($\Omega^{-1}cm^{-1}$)	Diffusion Coefficient (cm^2/sec)	R
	High-Purity AgBr		
110·08	$8{\cdot}32\times10^{-6}$		
123·17	$1{\cdot}94\times10^{-5}$		
140·64	$5{\cdot}24\times10^{-5}$		
145·39		$4{\cdot}26\times10^{-10}$	2·07
157·75	$1{\cdot}18\times10^{-4}$		
174·24	$2{\cdot}37\times10^{-4}$		
180·25	$2{\cdot}97\times10^{-4}$		
181·45		$1{\cdot}88\times10^{-9}$	1·97
197·56	$5{\cdot}83\times10^{-4}$		
210·14		$6{\cdot}62\times10^{-9}$	1·78
218·70	$1{\cdot}27\times10^{-3}$		
245·21	$3{\cdot}20\times10^{-3}$		
277·08	$9{\cdot}25\times10^{-3}$		
	Reagent-Grade AgBr		
126·55	$1{\cdot}83\times10^{-5}$		
143·36	$3{\cdot}26\times10^{-5}$		
149·35		$3{\cdot}40\times10^{-10}$	1·33
157·51	$6{\cdot}14\times10^{-5}$		
164·38		$6{\cdot}45\times10^{-10}$	1·49
173·78	$1{\cdot}41\times10^{-4}$		
186·37		$2{\cdot}00\times10^{-9}$	1·63
190·80	$3{\cdot}34\times10^{-4}$		
204·35		$4{\cdot}81\times10^{-9}$	1·60
206·38	$6{\cdot}85\times10^{-4}$		
223·66	$1{\cdot}40\times10^{-3}$		
227·35		$1{\cdot}35\times10^{-8}$	1·54
242·11	$2{\cdot}77\times10^{-3}$		
248·34		$2{\cdot}98\times10^{-8}$	1·55
261·00	$5{\cdot}33\times10^{-3}$		
270·37		$6{\cdot}82\times10^{-8}$	1·50
282·81	$1{\cdot}10\times10^{-2}$		
296·36		$1{\cdot}64\times10^{-7}$	1·49
304·87	$2{\cdot}17\times10^{-2}$		
333·90	$5{\cdot}15\times10^{-2}$		

FRIAUF[2] are included for comparison. Our conductivities are approximately 10 per cent smaller than those of FRIAUF and our diffusion coefficients are from 3 to 10 per cent smaller than his. Table 1 presents the conductivity and diffusion-coefficient data. The measured diffusion coefficients are also shown in Fig. 3. The quantity

$$R = \left(\frac{\sigma}{D}\right)\left(\frac{kT}{Ne^2}\right) \qquad (2)$$

VOL. 4 1958

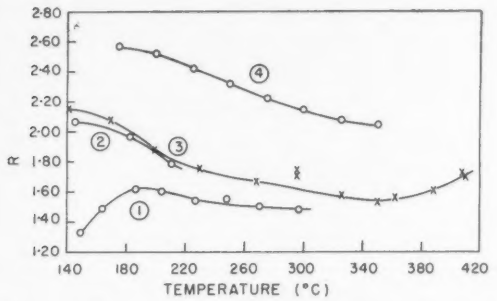


FIG. 4. The ratio $R = (\sigma/D)(kT/Ne^2)$ vs. temperature. (1) Reagent-grade silver bromide. (2) High-purity silver bromide. (3) Data of FRIAUF.[3] (4) Calculated from data of TELTOW[6] assuming collinear interstitialcy jumps.

is listed in Table 1 and shown in Fig. 4 as a function of temperature. R is unity if the Einstein relation is valid. FRIAUF's results are also shown in Fig. 4, as well as a plot of the relationship:

$$R = 3\left[\frac{\phi+1}{\phi+2\cdot4}\right] \qquad (3)$$

where ϕ is the ratio of the mobility of the interstitial silver ion to the mobility of the silver-ion vacancy, as determined by TELTOW.[6] Equation (3) yields the value of R which is predicted if Frenkel disorder of the silver-ion lattice is assumed and if it is assumed that the interstitial silver-ions migrate by collinear interstitialcy jumps. FRIAUF has shown that the assumption that a fraction of the interstitialcy jumps are noncollinear permits R, as calculated from TELTOW's mobility data, to be brought into agreement with the value of this quantity obtained from equation (2). In equation (2), kT is the product of Boltzmann's constant and the absolute temperature. The ionic charge is e. $N = 2\cdot08\times10^{22}$ cm^{-3} is the number of silver-ion lattice sites per unit volume. The change of N with temperature which results from the thermal expansion of the lattice has been neglected. The dimensions of the conductivity samples were measured at room temperature, and the sectioning of the diffusion samples was performed at room temperatures. Temperature corrections which should be applied because of these procedures have also been neglected, because the resultant change in R is comparable to the experimental error.

Fig. 5 shows the observed diffusion coefficient and that calculated from the measured conductivity by the use of equation (1) for a silver bromide crystal containing 0·064 mole per cent of cadmium

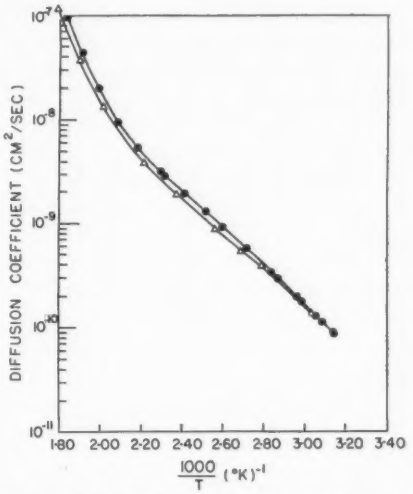


FIG. 5. Self-diffusion coefficient of silver in silver bromide containing 0·064 mole per cent cadmium bromide. △ Measured. ● Calculated from conductivity with $R = 1$.

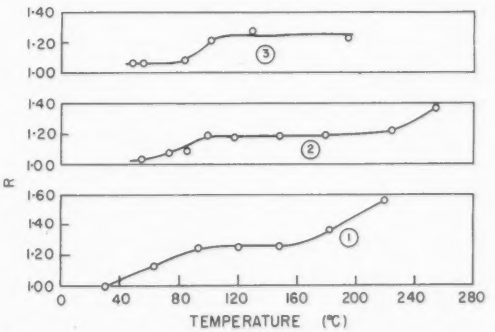


FIG. 6. The ratio $R = (\sigma/D)(kT/Ne^2)$ vs. temperature. (1) 0·028 mole per cent cadmium; (2) 0·064 mole per cent cadmium; (3) 0·075 mole per cent cadmium.

bromide. Fig. 6 shows the quantity R for three silver bromide–cadmium bromide crystals as a function of temperature.* There exists a well-

* The detailed conductivity and diffusion data in tabular form are available in a report to the Air Force Office of Scientific Research AFOSR-TN-S7-224, Astia Document No. AD. 126–522.

defined region of temperature for each of these crystals within which R has the value 1·25 to within the experimental error. This is the value of R predicted by the Bardeen–Herring correlation factor for vacancy diffusion in a face-centered lattice. We conclude that the addition of the divalent cadmium ion to the crystal has introduced a sufficient concentration of silver-ion vacancies to reduce to insignificance the concentration of interstitial silver ions. Migration of the silver ion occurs exclusively by the vacancy mechanism in the absence of interstitial silver-ions. At high temperatures, R rises as the concentration of interstitial silver-ions becomes significant. As expected, the larger the cadmium content, the higher the temperature at which this rise begins. At low temperatures, R falls below the value of 1·25. This drop in R is undoubtedly due to the association of cadmium ions with positive-ion vacancies to form complexes and larger aggregates.

Acknowledgements—One of the authors (A.S.M.) wishes to thank the Sylvania Electrical Products, Inc., for the award of a Fellowship.

The determination of the cadmium concentration of the crystals was performed at the Naval Research Laboratory through the co-operation of Dr. JAMES SCHULMAN. The assistance of Dr. W. DALE COMPTON is gratefully acknowledged.

REFERENCES

1. COMPTON W. D. *Phys. Rev.* **101,** 1209 (1956).
   COMPTON W. D. and MAURER R. J. *J. Phys. Chem. Solids* **1,** 191 (1956).
2. FRIAUF R. J. *Phys. Rev.* **105,** 843 (1957).
3. TANNHAUSER D. Thesis, University of Chicago (1957).
4. MCCOMBIE C. W. and LIDIARD A. B. *Phys. Rev.* **101,** 1210 (1956).
5. KOCH E. and WAGNER C. *Z. phys. Chem.* B **38,** 295 (1937–38).
6. TELTOW J. *Ann. Phys., Lpz.* **5,** 63, 71 (1949).
7. SEITZ F. *Acta Cryst.* **3,** 355 (1950).
8. BARDEEN J. and HERRING C. *Imperfections in Nearly Perfect Crystals* (Edited by W. SHOCKLEY). John Wiley, New York (1952).
9. COMPAAN K. and HAVEN Y. *Trans. Faraday Soc.* **52,** 786 (1956).
10. LIDIARD A. B. *Handbuch der Physik* Vol. 20. Springer-Verlag, Berlin (1956).
11. EBERT I. and TELTOW J. *Ann. Phys., Lpz.* **15,** 268 (1955).
12. HOVE J. *Phys. Rev.* **102,** 915 (1956).
13. MAPOTHER D., CROOKS H. N. and MAURER R. J. *J. Chem. Phys.* **18,** 1231 (1950).
14. LUCKEY G. and WEST W. *J. Chem. Phys.* **24,** 879 (1956).

VOL. 4 1958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 201–205.

INFRARED ABSORPTION BY SODIUM CHLORIDE-TYPE LATTICES

HERBERT B. ROSENSTOCK

U.S. Naval Research Laboratory, Washington, D.C.

(*Received* 19 *August* 1957)

Abstract—A computation is made of the frequencies at which the vibrational motion of the atoms in a crystal of the NaCl-type structure should produce absorption. Four subsidiary frequencies are found in addition to the classical "limiting frequency". The results are evaluated for those alkali halides for which compressibility and lattice distance are known; it is found that experimental absorption measurements made on thin evaporated films are fitted more closely by some of the subsidiary frequencies than by the classical limiting frequency.

INTRODUCTION

In two recent papers[1,2] we have attempted to calculate some features of the infrared spectrum of ionic crystals due to the motion of the atoms.* It was shown that if the lattice-dynamical problem is solved with the correct boundary conditions, one main absorption peak and several more subsidiary peaks are found, in contrast to the usual result of one simple absorption line which is obtained by simplifying the calculation by assuming "cyclic" boundary conditions.[3] The additional peaks were shown to arise from those points in K-space at which a critical point of the frequency surfaces coincides with maximum wavelength in one direction—points which, for cubic crystals, occur only at the corner of the periodicity cube in K-space.

A rigorous calculation of the lattice absorption spectra of a crystal would involve the calculation of the probability for absorption of electromagnetic radiation of its resonant frequency by every normal mode of the crystal, or a calculation of the classical dipole moment of the mode, to which the absorption probability is related.[2] The practical problem is simplified considerably, however, once it is shown that the absorption is essentially negligible at all but a finite number of frequencies; only (a) the frequencies at which the absorption probability is not negligible and (b) the absorption probability at those frequencies need be found. In I the rigorous problem was solved for a one-dimensional structure, and qualitative conclusions were drawn for three-dimensional crystals. In II the simplified problems (a) and (b) were both solved for a three-dimensional model of an ionic crystal which had the virtue of being both nontrivial for our purposes and mathematically tractable, but had the drawback that the forces assumed were not physically realistic.

The purpose of the present paper is the solution of problem (a), the position of the absorption lines, for crystals of the sodium chloride structure. This is possible in a straightforward manner—we merely have to calculate the frequencies at the corners of the periodicity cube in K and the results will provide that information about the spectrum that is most likely to be useful in comparison with experiment. We have not attempted here to solve problem (b), the actual intensities, because of the combination of much greater computational difficulty (the exact form of the normal modes would have to be known) with lesser importance to experiment.

CALCULATIONS

We calculate the optically active frequencies of the NaCl-type lattice by setting up the equation of motion for the atoms, deriving the secular equation for the frequencies and solving it at only those points in reciprocal space that correspond to

* These two papers will be referred to below as I and II, respectively.

adjacent atoms being (nearly) completely out of phase. In the monatomic model treated in II, these points were located at the points $(\pi\pi\pi)$; $(\pi\pi 0)$, $(\pi 0\pi)$, $(0\pi\pi)$; and $(\pi 00)$; $(0\pi 0)$, (00π). In the case of the present diatomic lattice, the position of the points depends on the way in which wave numbers are defined, which in turn depends on the choice of the unit cell, which is arbitrary. The unit cell obtained by choosing, in the usual way, two atoms per cell is not a cube, and the resulting assymmetry complicates both the locating of the optically active points and the form of the coulomb sums which must be evaluated; we therefore choose a cubic unit cell containing eight atoms (Fig. 1).

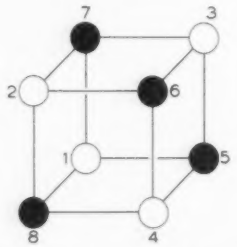


FIG. 1. The atoms in one unit cell. White atoms have charge e, mass m_1; black atoms have charge $-e$, mass m_2. The distance between nearest neighbors is a.

With that choice we have to solve the secular equation only at the one point (000) for which all cells are in phase. To be sure, this choice leads to a 24×24 determinant instead of the 6×6 determinant obtained from a choice of two atoms per unit cell; but at the point (000) this determinant is quite tractable.

The potential energy of interaction between two atoms in the lattice is taken as

$$\Phi = \Phi^{\text{rep}} + \Phi^{\text{coul}} \tag{1}$$

and the potential energy of the entire lattice is the sum of (1) over all particle pairs. The coulomb interaction for two particles whose equilibrium positions are ka, la, ma, respectively, apart in the x, y, and z directions (where a is the distance between nearest neighbors) is

$$\Phi^{\text{coul}} = (-)^{k+l}e^2/r \tag{2}$$

with r the distance between them. The repulsive interaction, being short-range, may for our purposes be neglected for all but nearest neighbors and may be expanded in the form

$$\Phi^{\text{rep}} = \Sigma\, \xi_l (r-\alpha)^l \tag{3}$$

where the constant ξ_1 can be determined from the fact that the lattice is in equilibrium when $r = a$, and ξ_2 from the known compressibility R[4,5], and all other ξ_l maybe neglected. We find

$$\xi_1 = e^2M/6a^2, \qquad \xi_2 = 3a/2R + e^2M/6a^3 \tag{4}$$

where $M = 1{\cdot}74755$ is the Madelung constant for the NaCl lattice.

Next the potential is expanded in terms of the displacement u_{klm}, v_{klm}, w_{klm}, of the particles from equilibrium in the x, y, z directions, respectively, keeping only quadratic terms and the x component of the force on particle k, l, m is found by differentiating the potential with respect to u_{klm}; the normal mode transformation is then applied to the displacements,

$$\left.\begin{matrix} u_{klm} \\ v_{klm} \\ w_{klm} \end{matrix}\right\} = \exp i[\omega t + 2kp + 2lq + 2mr] \left\{\begin{matrix} U_j \\ V_j \\ W_j \end{matrix}\right. . \tag{5}$$

U, V, and W are constants, the subscript j runs from one to eight according to Fig. 1, (p, q, r) are phase or momentum components, and ω is the frequency. This is then put into the Newtonian equations of motion, and the secular determinant consisting of the coefficients of the U_j, V_j, W_j is obtained as the consistency condition for them; it may then be solved for the frequencies.

The factor 2 appears in equation (5) because we have chosen the integers k, l, m, to label atoms rather than cells of 2^3 atoms; this is a trivial point, since we shall restrict ourselves to $(p, q, r) = (0, 0, 0)$ anyway. Less obviously trivial is the fact that the normal modes (5) have been chosen to be exactly of the form which is consistent with cyclic boundary conditions and which, according to I, will therefore fail to give the subsidiary absorption lines, the location of which is the main purpose of the present work. The correct modes differ from those in (5) by a small but, for the present problem, unknown phase factor. The choice (5) is nevertheless satisfactory; for it is only the actual calculation of the absorption probabilities that is sensitive to small phase changes. The calculation of the frequencies, which is all we require here, is essentially independent of the relative phases.

The phases (p, q, r) are then equated to $(0, 0, 0)$. Those coefficients of the secular determinant which correspond

to force in the x-direction due to motion in the y- or z-direction are then seen to vanish (the repulsive part of them identically, the coulomb part because contributions from atoms $(k; l, m)$ and $(k; -l, -m)$ cancel). As a result the 24×24 determinant factors into three identical 8×8 determinants which may be written

$$|-\mathcal{J}-\mathcal{M}w^2|=0 \tag{6}$$

where

$$\mathcal{M}_{ij}=\delta_{ij}\begin{cases}m_1 & 1\leqslant i\leqslant 4\\ m_2 & 5\leqslant i\leqslant 8\end{cases} \tag{7}$$

and the elements of $\mathcal{J}$ contain ξ_1 and ξ_2 as contributions from the repulsive interaction and as a contribution from the coulomb interaction eight different sums of the form

$$\sum_{\substack{k,l,m\\-\infty}}^{\infty} f(k;l,m)$$

$$=\sum_{\substack{k,l,m\\-\infty}}^{\infty}(-)^{k+l+m+1}(2k^2-l^2-m^2)/(k^2+l^2+m^2)^{5/2} \tag{8}$$

taken over different sets of lattice points $(k; l, m)$. The symmetry of $f(k; l, m)$ can be used to show that of these lattice sums only two are both nonvanishing and different. One finally gets

$$-\mathcal{J}=\begin{bmatrix}-\mathcal{J}_1^* & -\mathcal{J}_2^*\\ -\mathcal{J}_2^* & -\mathcal{J}_1^*\end{bmatrix} \tag{9}$$

with the 4×4 submatrices $\mathcal{J}^*$ given by

$$-\mathcal{J}_1^*=\beta\begin{bmatrix}\rho & W & -W/2 & -W/2\\ W & \rho & -W/2 & -W/2\\ -W/2 & -W/2 & \rho & W\\ -W/2 & -W/2 & W & \rho\end{bmatrix}$$

$$-\mathcal{J}_2^*=\beta\begin{bmatrix}V'-\rho & 0 & -V'/2 & -V'/2\\ 0 & V'-\rho & -V'/2 & -V'/2\\ -V'/2 & -V'/2 & V'-\rho & 0\\ -V'/2 & -V'/2 & 0 & V'-\rho\end{bmatrix}$$

with

$$\beta=e^2/a^3.\ \rho=6a^4/Re^2.\ V'=V-2M/3 \tag{11}$$

and with V and W sums (8) taken, respectively, over points $(k; l, m)$ = (odd; even, even) and (even; odd, odd). Equation (6) can be solved in the usual way by adding and subtracting columns and rows. We find

$$\omega_1=0$$

$$\mu\omega_2^2/\beta=\rho$$

$$2\mu\omega_3^2{}_\pm/\beta=A_1\pm\sqrt{[A_1^2-(4\mu^2/m_1m_2)(A_1^2-A_2^2)]}$$

$$2\mu\omega_4^2{}_\pm/\beta=B_1\pm\sqrt{[B_1^2-(4\mu^2/m_1m_2)(B_1^2-B_2^2)]}. \tag{12}$$

Here

$$\mu^{-1}=m_1^{-1}+m_2^{-1}$$

$$A_1=\rho-W \qquad A_2=-\rho+V'$$

$$B_1=\rho+2W \qquad B_2=-\rho+2V'.$$

The two solutions $\omega_{3\pm}{}^2$ each occur doubly.

Of the eight solutions (12), ω_1 is the trivial translation. ω_2 is the so called "limiting frequency"—the only optically active frequency according to the old theory in which correct boundary conditions are not taken into account; this one could have been obtained in a much simpler way.[6] The six others are the ones at which the "subsidiary" peaks that particularly interest us occur. For a large crystal ($N\to\infty$), ω_2 would have the greatest intensity, ω_{3+} and ω_{4+} the next greatest and ω_{3-} and ω_{4-} the next greatest; for small crystals (roughly defined as crystals in which a substantial fraction of the atoms are close to one of the surfaces), these relative intensities may, of course, be reversed or otherwise changed.

RESULTS

Equation (12) for ω contains two kinds of quantities. The ones denoted by lower-case Greek letters are parameters involving the atomic masses, the nearest-neighbor distances, and the compressibilities of the individual compounds; the ones denoted by capital letters are lattice sums which depend only on the lattice structure.

The former quantities were looked up;[7,8] the average atomic weights for the usual isotopic composition was used—a point not likely to cause large errors, but worth mentioning because it does violate our assumption that only two masses occur in the crystal.

Of the lattice sums, $M = 1{\cdot}7455$ is known[9] and W and V were evaluated by a mathematically elementary method due to FRANK,[9] slightly modified. The method may be described as follows: sum the series directly inside a cubic region K centered about the origin. Outside that region divide space up into cubes such that the atoms appear only at their centers, and then replace the actual distribution by a sum of two distributions, $a+b$, equal to the original; a is obtained by dividing each atom into 8 equal parts and putting $\frac{1}{8}$ at each corner of its cube, while b is obtained by leaving the atoms put at the center and in addition placing the negative of $\frac{1}{8}$ atom at each corner of its cube. If the terms in the sum being evaluated alternate in sign, contributions to a appear only

on the surface of K, where they are readily evaluated, and the sum b may be neglected, being a multipole system of several orders higher than the original sum. Since the terms of our sum W do not alternate in signs, the actual sum evaluated was not $W = \sum_{\text{even; odd, odd}} f(k; l, m)$, but $W - \sum_{\text{even; even, even}} f(k; l, m)$, which is equal to W because the second sum vanishes on account of the symmetry properties of f; and similarly for V. The speed of conversion of FRANK's method may be illustrated by quoting the successive results for W obtained by letting the edge of K be respectively 3, 5, 7, 9: they are 1·01405, 1·08028, 1·08360, 1·08248. The final results used were $W = 1{\cdot}083$, $V = 3{\cdot}754$.

The results for all alkali halides of the NaCl structure for which compressibility data were available are presented in Table 1. We note that most of the observed "subsidiary" wavelength—and in some cases all of them—are larger than λ_2.

Table 1. *Wavelengths absorbed by alkali halides, as computed from equation* (12)

$\lambda_{i\pm} = \pi c/\omega_{i\pm}$, in microns

	λ_2	λ_{3+}	λ_{4+}	λ_{3-}	λ_{4-}
LiF	19·5	28·8	16·2	49·2	34·1
LiCl	28·8	33·7	22·4	86·3	52·6
LiBr	31·0	39·3	25·6	133·5	91·6
LiI	35·9	44·1	29·1	206·5	130·1
NaF	30·5	43·6	29·8	66·9	40·8
NaCl	45·6	60·2	46·2	107·0	66·7
NaBr	54·6	69·1	49·8	162·6	95·5
NaI	64·6	80·0	56·3	227·6	136·6
KF	39·8	52·4	38·6	101·4	58·0
KCl	57·2	74·0	63·4	141·1	70·3
KBr	72·5	91·6	72·7	195·0	105·0
KI	84·4	103·3	79·9	260·5	144·5
RbCl	70·7	89·0	68·9	196·5	108·4
RbBr	97·2	125·0	108·1	240·5	118·6
RbI	114·5	143·6	123·4	301·5	150·9
CsF	48·0	56·2	42·7	200·1	117·0

DISCUSSION

(a) *Comparison with other theoretical work*

In the infinite crystal the frequency ω_2 results from the very simple mode in which all like atoms are in phase and are all moving in the same direction towards one of their nearest neighbors. This frequency can therefore be calculated very simply[6] and our results agree rather well with those calculations.

Detailed calculations of frequency spectra have been made for NaCl by KELLERMAN[4] and for KCl (in the monatomic approximation) by IONA;[10] a less complete one for NaCl has been made by LYDDANE and HERZFELD.[11] Our results appear to be in fair agreement with theirs, if one remembers that IONA's input data (elastic constants) were adjusted to make his ω_2 fit the experimental "residual wave" frequency ($\lambda_2 = 70{\cdot}7$ microns). We do not obtain IONA's extra frequency at wave number (π, π, π) which has been criticized elsewhere[4,12] as physically unrealistic owing to improper treatment of infinite wavelength.

(b) *Comparison with experiment*

No measurements of the far infrared absorption by alkali halides seem to have been made since the 1930's.[13,14] The present series of papers was in part motivated by the fact that these results are not at all well described by one single frequency, as classical theory would demand.[14] Experiment gives absorption bands with a definite structure which seems to depend on sample thickness. In addition the classical calculation of the single absorption frequency (see (a) above) did not even agree well with the observed absorption peak, the discrepencies being of the order of 20 per cent.[6]

The most interesting result of the present paper is that in every case the observed absorption peak is matched much better by the one or more of "subsidiary frequencies", λ_{3+} in most cases, than by the classical "limiting frequency" λ_2. (See Table of reference 6 and Fig. 2 of reference 14). In view of the fact that the samples used were evaporated films at most a few microns in thickness, it should not be surprising if the relative intensities predicted for infinite crystals has been reversed.

On the other hand it must be realized that the samples, which were not ideal infinite crystals, could hardly have been ideal finite crystals either, and one can hardly expect the results of the present calculation to apply to them quantitatively. The arts of crystal growing and crystal cleaving, as well as of infrared spectroscopy, have made great progress in the last twenty years, and it may be hoped that someone will feel it worthwhile to make measurements in the future against which the present results may be checked.

VOL 4 195

REFERENCES

1. ROSENSTOCK H. B. *J. Chem. Phys.* **23,** 2415 (1955).
2. ROSENSTOCK H. B. *J. Chem. Phys.* In press.
3. BORN M. and MAYER M. G. *Handbuch der Physik* Vol. 24, Part 2, p. 652. Springer-Verlag, Berlin (1933).
4. KELLERMAN E. W. *Proc. Roy. Soc.* A **238,** 513 (1940).
5. KITTEL C. *Introduction to Solid-State Physics* (2nd Ed.) pp. 77–79. John Wiley, New York (1953).
6. KITTEL C. *Ibid.* pp. 112–114.
7. WYCKOFF R. W. G. *Crystal Structures* Vol. I, Table III 2. Interscience Publishers, New York (1948).
8. BIRCH F., SCHLAIRER J. F. and SPICER A. C. *Handbook of Physical Constants* p. 52. Geophysical Society of America (1942).
9. FRANK F. C. *Phil. Mag.* **41,** 1287 (1951).
10. IONA M. *Phys. Rev.* **60,** 822 (1941).
11. LYDDANE R. H. and HERZFELD K. F. *Phys. Rev.* **54,** 864 (1938).
12. BORN M. and HUANG K. *Dynamical Theory of Crystal Lattices* Chap. V. University Press, Oxford (1954).
13. BARNES R. B. *Z. Phys.* **75,** 723 (1932).
14. BARNES R. B. and CZERNY M. *Z. Phys.* **72,** (1931); Fig. 2 is reprinted in References 12, p. 125.

VOL.
4
1958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 206–216.

TRANSIENT PHOTOCONDUCTIVITY IN AgCl AT ROOM TEMPERATURE*

FREDERICK C. BROWN

University of Illinois, Urbana, Illinois

(*Received* 5 *August* 1957)

Abstract—At room temperature silver chloride is an ionic conductor having a relaxation time for polarization of several hundred μsec. A fast-pulse technique is described which makes possible the separation of electronic from ionic currents in this material at 25°C. Primary photoresponse in essential agreement with a simple trapping model is observed. The factors which influence electron lifetimes in pure silver chloride were investigated for crystals containing only a few tenths of a part per million of heavy-metal impurities. A relatively low photoconductive response corresponding to electron lifetimes before trapping of 10^{-8} sec or less was found for crystals grown from the melt in vacuum, in a halogen, or in an inert atmosphere. This applies to well-annealed samples and to materials for which precautions were taken to remove water and the products of hydrolysis (Ag and Ag_2O). On the other hand, when crystals are grown in air or from a moist starting charge, electron lifetimes as long as 10 μsec may result. The contribution of holes to fast photoresponse was not observed and must be quite small. For example, in the case of a crystal having a large response at 1000 V/cm field, the ratio of hole range to electron range was found to be $< 0{\cdot}012$, yielding an upper limit of hole lifetime of 4×10^{-9} sec. From measurements with single pulses, the drift mobility of electrons at 26°C is found to be 46 ± 4 cm²/V sec, in agreement with print-out Hall experiments.

VOL. 4 1958

INTRODUCTION

At low temperatures the dark conductivity of photoconducting AgCl is very small. In other words, if we regard a crystal of this material as a capacitor and resistor in parallel, the relaxation time, $T_R = RC$, is very much longer than the transit time of charge carriers. As explained by Van Heerden,[1] this is one requirement for the observation of primary photocurrents. As is well known, Pohl and his collaborators demonstrated the existence of these primary currents in the silver halides many years ago.[2,3] Lehfeldt,[3] for instance, measured saturation photocurrents, using a sensitive electrometer and low light intensities. His crystals, cooled to 77°K, were not the same before and after illumination, as evidenced by the fact that traps become filled and polarization effects gradually set in. Nevertheless, experiments of this type in their initial stages lend themselves to a relatively simple interpretation in terms of electron ranges governed by a volume distribution of traps. More recently, transit-time measurements of carriers have yielded values of electron mobility.[4]

At room temperature AgCl is an ionic conductor. Even for the purest crystals, as shown by C–38–5 and X_2 in Fig. 1, dark conductivity lies in the range 10^{-8} to 10^{-7} Ω^{-1}cm^{-1} at 300°K. If, at these temperatures, electrodes could be applied which permitted entry of a charge upon application of a collecting voltage, electrolytic decomposition would occur owing to ion transport. On the other hand, application of a voltage to blocking electrodes establishes an electric field which rapidly relaxes to zero in the interior of the dielectric. This occurs because of the motion of ionic effects (principally positive-ion vacancies in AgCl at 300°K) and the establishment of a space charge near one or both surfaces of the dielectric.

A theory for steady-state photocurrents in the presence of dark current has been given by Stöckmann.[5] His results apply either to the case of electronic semiconductors or to colored alkali halides at high temperatures, as studied by Hilsch and Pohl.[6] In the present paper, we deal

* Partially supported by the U.S. Air Force.

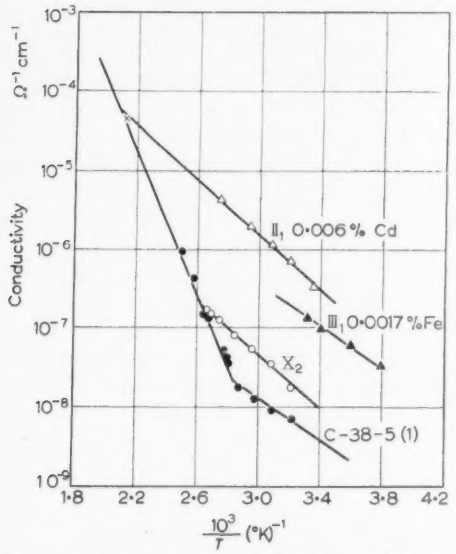


FIG. 1. Ionic conductivity as a function of $1/T$ measured by the a.c. bridge technique.

with transient rather than stead-state photoconductivity in the presence of mobile ionic defects. Using fast-pulse techniques, it is possible to separate electronic from ionic currents and to observe a primary photoresponse which is in essential agreement with a simple trapping model.

POLARIZATION DUE TO ION DRIFT

A complete analysis of the polarization problem for blocking electrodes involves the solution of a rather complex nonlinear differential equation.[7] It is of use for the case at hand to take a less general point of view and to calculate the fields which arise during the *onset* of polarization. Let us find a relation for the electric field within the volume of an unilluminated crystal a time t after application of a step-voltage waveform of amplitude V, as shown schematically in Fig. 2 by closing the switch. Assume plane parallel blocking electrodes, a crystal thickness, d, and initially a uniform density of *negative* charge, ρ_0, mobile throughout the volume. This is the case for positive-ion vacancies, the main ionic defect, and, incidentally, also the case for uniform volume illumination of a photoconductor in which electrons are mobile and holes are fixed. After a time t, the negative charge will have displaced an average distance, w_i, giving rise

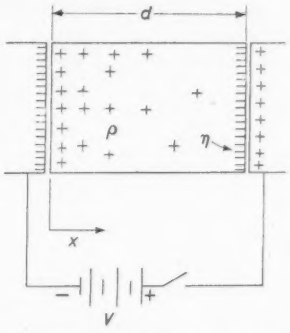


FIG. 2. Distribution of net charge, ρ, in a crystal due to polarization t sec after application of a voltage, V. Initially it is assumed that the crystal is neutral throughout its volume, but that a uniform density of negative charge, ρ_0, is free to move.

to a total charge density in the volume of the crystal

$$\rho = \rho_0 e^{-x/w}. \tag{1}$$

We make the simplifying assumption that w_i is constant through the crystal, an assumption valid if the applied field is large compared to the polarization field. There will also be an appreciable surface density of charge adjacent to the positive electrode (blocking electrodes) given, per cm^2, by

$$\eta = \rho_0 w_i (1 - e^{-d/w_i}). \tag{2}$$

A total charge will flow through the external circuit during the time that polarization sets in, given by[8]

$$q = \rho_0 d \left\{ \frac{w_i}{d} \left[1 - \frac{w_i}{d} (1 - e^{-d/w_i}) \right] \right\} \tag{3}$$

if the resistance of the external circuit is not too great. These charges, as well as the charge on the electrodes when the voltage V is applied, are all sources of displacement, D. A straightforward application of Gauss' law, erecting a pillbox having one surface within the crystal, the other at $-\infty$, yields the electric field $E(x)$ in e.s.u. as follows:

$$E(x) = \frac{4\pi\rho_0 w_i}{K} \left[\frac{w_i}{d} (1 - e^{-d/w_i}) - e^{-x/w_i} \right] - \frac{V}{d}. \tag{4}$$

The last term, $-V/d$, is the applied field in the absence of polarization and is, of course, opposite in direction to the first terms (assumed small)

which arise because of ion drift. K is the dielectric constant of the material. In the limit that $w_i \ll d$, valid for ionic conductivity at the onset of polarization:

$$E \simeq \frac{4\pi\rho_0 w_i^2}{Kd} - \frac{V}{d} \qquad w_i \ll d \qquad (5)$$

except when $x \ll d$.

Note that, except very near the cathode, the field is uniform in the volume of the crystal, a consequence of the assumption that ion drift is small. Let us make the substitutions $\rho_0 w_i = \sigma(V/d)t$ and $w_i = \mu_i(V/d)t$, where σ is the conductivity of the crystal (ρ_0 = initial mobile charge density).

Then we have the result

$$E \simeq \left(\frac{4\pi\sigma\mu_i V t^2}{Kd^2} - 1\right)\frac{V}{d}. \qquad (6)$$

The mobility of positive-ion vacancies, μ_i, at room temperatures in AgCl is not known accurately. However, the data of KOCH and WAGNER[9] indicate it may be as high as 10^{-5} cm sec/V cm. For $\sigma = 3\times10^{-8}\ \Omega^{-1}\text{cm}^{-1}$, a crystal of 0·5 cm, and fields as high as 10^3 V/cm, the first term in equation (6) begins to approach 1 per cent only for times, t, in the vicinity of several msec. Just as at low temperature, the relaxation time is longer than the expected transit time of charge carriers. Furthermore, in the early stages of polarization under the assumptions stated above, the field is reasonably independent of x.

Consideration of the fluctuations of ion current in AgCl at room temperature indicates that they will not mask a small ionization or photocurrent. These results and the estimates made above make feasible the separation of the fast drift of electrons from the very much slower displacement of ions under the influence of a pulsed field. The technique will be described in the next section.

EXPERIMENTAL METHOD

Single flat-top voltage pulses, balanced to ground, having a rise time of about 10 μsec and a duration of 1·0 msec,* are applied to two identical crystals mounted in each arm of the high-frequency bridge circuit shown schematically in Fig. 3. The crystals, of thickness 0·23 cm, rest on transparent semiconducting glass plates, as shown in Fig. 4 of the electrode arrangement. Thin sheets of Mylar, 0·00025 in. thick, separate the crystal from the electrodes and assure blocking conditions. (At voltages above 400 V, this had the effect of reducing noise due to the charge injected from the

VOL. 4 1958

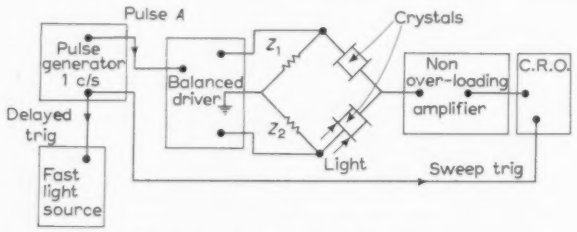
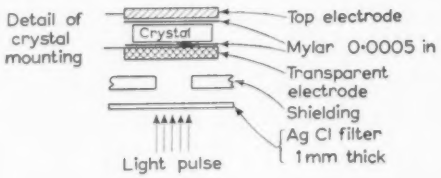


FIG. 3. Block diagram of apparatus.



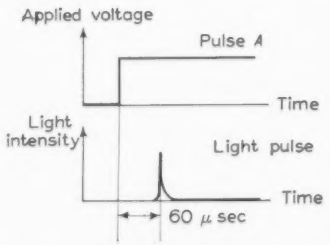


FIG. 4. Electrode arrangement and pulse sequence.

* Supplied by a Tektronix type-162 waveform generator followed by a paraphase driver and a push–pull amplifier employing two 6BG6s. Pulse amplitude = 0–1000 V either side of ground.

electrodes; at lower voltages, the Mylar could be omitted.)

In operation the bridge is carefully balanced without illumination so that the applied voltage pulses do not block the amplifying equipment. After application of a field pulse, one crystal is illuminated with a short (0·5 μsec) flash of light, while the other is kept in the dark. The bridge unbalance is detected by a high-gain wide-band system employing a direct-coupled pre-amplifier, a nonoverloading amplifier,[10] and a fast oscilloscope (Tektronix type 511). Very careful balance of the bridge must be achieved in addition to the good overload characteristics of the amplifier, since small photovoltages of the order of 100 μV must be detected shortly after the application of a collecting voltage of several hundred volts. The bridge circuit is symmetrical and of the equal-arm capacitance type. Thought was given to the impedance of its elements so that balance could be achieved for crystals having widely different characteristics. The variable resistance in shunt with the crystal was of the order of 100 kΩ and the capacitance of leads and pre-amplifier input kept to about 40 $\mu\mu$F. All signals were measured by comparing them (under operating conditions) to a small charge applied through a calibrating capacitance at the input.

The fast-pulsed light source could be triggered with a predetermined delay of a few to several hundred μsec. It consisted of two pointed, unenclosed molybdenum electrodes spaced about 1 mm apart and supplied at 5000 V in a conventional negative pulser circuit, employing a 3C45 hydrogen thyratron. Since an air spark of this type was found to be rich in ultraviolet, a quartz lens was used to focus light onto the crystal. The duration, as well as the intensity, of the emitted light depends upon the storage capacitor used. With a capacitance of 0·002 μF, the total duration was of the order of 0·1 μsec. More often 0·01 μF was used, which resulted in a light intensity as a function of time, as shown in Fig. 5.

Two conditions of operation of the detector were utilized, depending upon the value of the resistance from grid to ground at the input of the pre-amplifier. When this resistance was of the order of 1·0 MΩ or larger, the most sensitive condition prevailed, and charge, of the order of 10^{-14} C per pulse, could be detected. Under these conditions

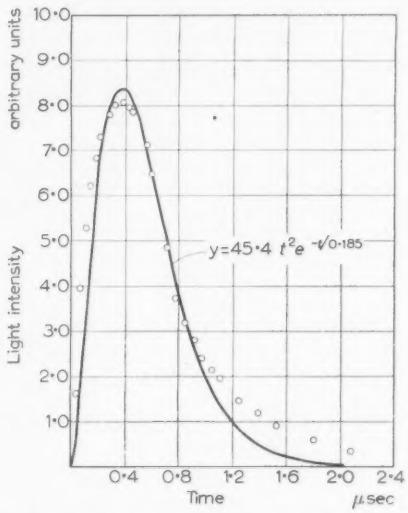


FIG. 5. Spark light intensity as a function of time. Storage capacitor 0·01 μF.

low-voltage pulses (0–60 V) were employed, and tests were carried out for the purpose of comparing the photoresponse of various samples. For light strongly absorbed near the surface, it was found that considerable variability existed from crystal to crystal owing to defective surface layers. These can be easily removed by prolonged polishing and etching, but for the comparison tests release of charge throughout the volume of the crystals was assured by filtering the light through 1·0 mm of AgCl. In the case of a range for electrons, w_e, and for holes, w_h, both small compared to the crystal thickness, d, the induced charge for N_0 excitations in the volume is given by:

$$Q \simeq \frac{N_{0e}}{d}(w_e + w_h). \tag{7}$$

The observed response is proportional to the range of carriers, $w_e = E\mu_e T_e$, where $E = V/d$ the electric field, μ_e the electron mobility and T_e the lifetime before trapping—similarly for holes.

Fig. 6 shows the charge response as a function of collecting voltage for a crystal of standard thickness, 0·23 cm under both positive and negative polarity. A 1·0 mm AgCl filter was employed, and it is to be noted that the positive and negative responses are identical because of the penetration

of the light. Fig. 7 shows photoresponse as a function of delay of the light pulse, measured from the start of the rise of the field pulse. One may observe that the internal electric field begins to relax in the vicinity of 225 μsec. Polarization began

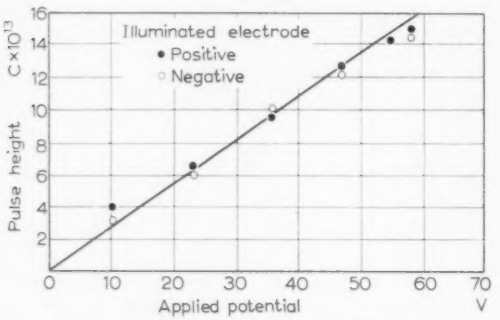


FIG. 6. Photocharge observed for single pulses of light which can penetrate the volume of the crystal. Crystal I_3, using a 1 mm AgCl light filter.

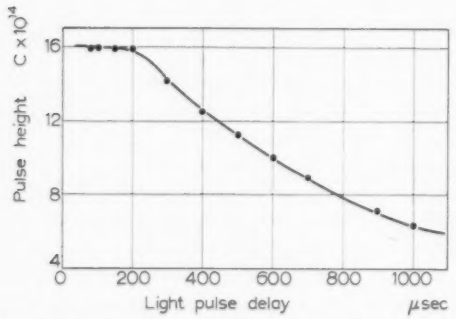


FIG. 7. Typical charge response versus light delay time at low voltage.

earlier in doped, higher-conductivity crystals, in thinner samples, or with higher applied voltages, quite in agreement with equation (6). For tests on average samples, it was satisfactory to set the delay in the vicinity of 50–100μsec.

In addition to the tests discussed above, the photoconductivity was studied as a function of light intensity over about two orders of magnitude, using neutral filters. The response was found to be linear with intensity under these conditions of low-level illumination—about 10^7 quanta absorbed per pulse. During the measurement, the polarization due to electron drift is small as well as that due to ion motion. The results of a measurement of charge Q for the case of penetrating light are given in the next section for a large number of samples of different preparation.

When the pre-amplifier input resistor is reduced to about 1000 Ω, current, $i(t)$, is measured, and it is possible to observe the decay of photoconductivity directly. For this method of operation, the AgCl filter is removed and an ultraviolet filter (Corning 5860, peak 360 mμ, absorption constant in AgCl at 26°K about 10^{+3}cm^{-1}) is substituted. The illuminated surface of the crystals must be carefully prepared to remove defective layers. It will be shown in a later section that the decay of photoconductivity in certain samples having long electron lifetimes is governed by an approximately uniform distribution of deep traps or, when high collecting voltages (up to 900 V) are used, by the drift of electrons across the thickness of the samples and collection at the upper electrode.

COMPARISON OF PHOTORESPONSE

(a) *Crystal preparation*

Control of the electronic properties of AgCl has been achieved by growing crystals in high vacuum, in chlorine, and in air, using the Bridgman technique. The main effort was directed towards producing single crystals of high purity, free from silver and silver oxide. A cross-section of the growing furnace is shown in Fig. 8. Starting material is prepared by precipitation in dilute, 0·1 N solutions of Kodak special product, X-491, silver nitrate, and redistilled Du Pont reagent hydrochloric acid. Vigorous mechanical stirring of the solutions is carried out, and the precipitation is stopped with the solution still excess in the chloride ion. The precipitate

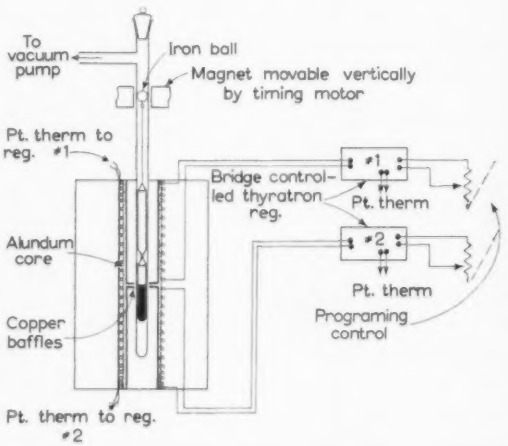


FIG. 8. Vacuum crystal-growing furnace. Lowering rates less than 1·0 mm/hr can be maintained.

Table 1. *Crystal properties*

Boule no.	Growth rate (mm/hr)	Remarks*	Sample no.	Charge response† ($C \times 10^{14}$)	
I	2·7	N.F.; wet precipitate	I_4	56	$T = 0{\cdot}35$ μsec
			I_5	142	$T = 1{\cdot}0$ μsec
II	2·7	N.F.; 0·006 per cent Cd added	II_1	1·0	
III	2·7	N.F.; 0·0017 per cent Fe added	III_2	9·0	
IV	1·0	F.	IV_1	3·6	
V	1·0	F.; raw material no. IV	V_1	1·0	
VI	1·0	F.	VI_1 Bot.	0·8	
			VI_2 Top	4·0	
VII	1·0	F.; platinum crucible	VII_1	1·8	
VIII	1·0	F.; platinum crucible raw material no. VII	$VIII_1$	1·8	
IX	1·0	F.	IX_1 Bot.	2·5	
			IX_2 Top	14	
X	1·0	F.	X_1 Bot.	1·0	
			X_2 Top	1·0	
XI	2·7	F.	XI_1 Bot.	1·1	
			XI_2 Top	4·5	
XII	2·7	F.; 0·01 per cent Ag_2O added	XII_1 Bot.	1·4	
			XII_2 Top	0·7	
XIII	2·7	F.; 0·035 per cent $AgNO_3$	$XIII_1$	1·1	
XIV	2·7	F.; moist O_2 added large excess	XIV_1	0·7	
XV	2·7	F.; 0·00012 per cent NaOH	XV_1	3·2	
XVI	2·7	F.; crucible lined with platinum	XVI_1	1·1	
XVII	2·7	N.F.; very dry precipitate	$XVII_1$ Top	12·6	
XVIII	2·7	F.; grown in air	$XVIII_1$ Bot.	80	
			$XVIII_3$ Top	340	$T = 10$ μsec
XIX	2·7	F.; ultra-high vacuum	XIX_1 Top	1·8	
XX	2·7	F.; Chlorine atmosphere $P = 4{\cdot}6$ mm Hg	XX_1 Bot.	1·4	
			XX_2 Top	1·4	
H–3		Harshaw, cut from ingot	H–3–2	142	$T \simeq 1{\cdot}0$ μsec
C–38–5		Kodak, helium atmosphere‡		3·6	
C–16–H721		Kodak, nitrogen atmosphere		2·8	
C–44–5		Kodak, nitrogen atmosphere		3·6	
H–342–196b		Kodak, Kyropoulos grown in air§	1	142	$T \simeq 1{\cdot}0$ μsec
			2	113	
G–649–36–1		Kodak, Kyropoulos grown in air	1	85	
C–63–3		Kodak, bubbled briefly with wet oxygen	1 Top	113	$T \simeq 1{\cdot}0$ μsec
C–63–6		Kodak, bubbled briefly with wet oxygen	1 Bot.	200	$T \simeq 2$ μsec

* F. = filtered through quartz capillary.
N.F. = not filtered; precipitate placed directly in growing crucible.

† Using a 1 mm AgCl filter and comparable pulse light intensity in each case. Crystal thickness = $0{\cdot}23 \pm 0{\cdot}02$ cm; collecting voltage = 60 V.

‡ Kodak Bridgman crystals kindly supplied by Dr. F. URBACH, Eastman Kodak Laboratories.

§ Kyropoulos crystals kindly supplied by Dr. FRANK HAMM, Eastman Kodak Laboratories.

is decanted and washed several times with very dilute hydrochloric acid. The material is then dried for 24 hr in a vacuum freeze drier of special design, in which the powder is heated to not more than 150°C. This results in a small-size granular precipitate free from hard lumps.

The washed and dried precipitate is next placed in a quartz charging tube in the vacuum growing furnace (Fig. 8), so that upon first fusing the molten salt will strain through a capillary into the growing crucible,[11] a G.E. fused-quartz test-tube, 15×1·5 cm. As a result, a small amount of material having adhesive properties (AgCl containing oxides) remains above the capillary.

All operations are carried out under darkroom red light, and care is taken to eliminate contamination, particularly from oil vapors. In vacuum or in chlorine, the crystal does not stick to the crucible walls upon cooling. Bubbles sometimes form on the surface of the boule, but these can be minimized by heating the molten salt in vacuum ($< 10^{-6}$mm Hg) at 495°C for several hours before establishing the growing conditions and lowering the crucible.

(b) *Annealing the samples*

The crystals after being formed are slowly cooled to room temperature. Samples of standard (0·23 cm) thickness are cut carefully from the boule with a thin carbide-tipped circular saw, especially designed for a soft corrosive material. X-ray analysis by means of back-reflection Laue patterns indicated that the crystals as grown were usually single or contained two regions separated by large-angle grain boundaries. The surface of the samples after cutting showed diffuse diffraction patterns and were highly disturbed. However, sharp patterns are easily obtained by polishing or etching away the defective layers (about 0·05 mm thick), indicating that the interior of the samples is not severely distorted.

Because of the importance of strains as electron traps,[12] tests were carried out to determine the effect of annealing at 400°C in various atmospheres—vacuum, air, and chlorine. The results may be briefly stated as follows:

(1) The photoresponse of most samples cut from the various crystals improved little or not at all upon annealing in vacuum, air, *or chlorine*. Strains were not very important when the crystals were carefully handled. This result applies to crystals resting on an inert substance during the heat treatment, such as on pure powdered quartz.

(2) A few crystals showed improvement after annealing, indicating that strains can be important, especially in those cases (air grown) where the material adhered strongly to the crucible walls during growth. The best annealing procedure appears to be in air on quartz powder at 400°C for several hours. Rising and cooling rates should be kept below 20°/hr.

(3) When crystals rest on powdered MgO or CaO *in air* during the anneal, a chemical reaction appears to take place at the surface. A remarkable improvement in electron lifetime occurs within 0·4 mm of the surface of pure vacuum-grown crystals treated in this way. For example, samples cut from crystal **X** showed a photoresponse of 1·0 arbitrary units before annealing and also after heat treatment in air on quartz. However, when annealed in air on MgO powder for 39 hr, the response increased to 426 arbitrary units. This special effect has been observed many times. By lapping away surface layers and from a study of the shape of current pulses, one deduces that the electron trap density is greatly reduced within 0·2–0·4 mm of the surface adjacent to the MgO powder. A diffusion process appears to take place in which the reaction at the surface is related to the presence of air and moisture and to the chemical properties of such bases as MgO and CaO. Annealing in this manner was not used in the comparison tests reported in the next section, nor when it was desired to study the properties of a homogeneous crystal.

(c) *Results of the comparison tests*

Table 1 gives pertinent data on twenty of our Bridgman crystals, together with crystals prepared by the group at Eastman Kodak Co.[13] The last column shows the charge response proportional to carrier range, w, for samples of standard thickness 0·23 cm.* All the tests were made at a collecting voltage of 60 V, using a AgCl filter and comparable pulse light intensity in each case. Electron lifetime is also shown when it could be directly observed, as explained in the next section. Table 2 gives the results of a careful spectrochemical analysis of two

Table 2. Results of spectrochemical analysis†

Element	Detectable limit (p.p.m.)	I_5	X_3	C-38
Cd	0·5			
Cr	0·05			
Co	0·1			
Cu	0·05	0·11	0·10	
Au	1·0			
Fe	0·05	0·26	0·33	0·11
Pb	0·075	0·06	0·09	
Mn	0·05			
Ni	0·05	0·05	0·05	
Sn	0·05			
Zn	4			
Mg	0·05	0·03		

† The detection limit and observed impurity are expressed in parts per million by weight. No entry means the element was not detected.

* Results for annealing on MgO have not been included.

of our samples and one Kodak crystal.* The results of comparing photoconductivity of many annealed and unannealed crystals may be stated as follows:

(1) Pure AgCl grown in vacuum, chlorine, or an inert atmosphere and from material for which care has been taken to remove moisture and the products of hydrolysis, (Ag and Ag_2O), shows relatively low photoresponse (see for example crystals IV through XI, XVI, XVII, XIX and several Kodak samples). This low response cannot be explained in terms of a decrease in primary yield of carriers alone, but arises because of a short electron lifetime, usually 10^{-8} sec or less.

(2) Crystals grown *in air*, without filtering, or in vacuum from a moist starting charge, show a much larger photoconductivity and electron lifetimes of the order of μsec. (For example refer to numbers I, XVIII, H-3, C-63-3, C-63-6 and the Kodak air-grown Kyropoulos crystals.) In earlier work also several air-grown crystals of lower purity showed a large response.[4] Attempts to add oxygen deliberately in the form of $AgNO_3$ or NaOH have not been particularly informative.

The author suggests that the density of electron traps is decreased by the presence of minute traces of silver oxide dissolved in AgCl. It is not apparent that these traps arise from a lack of stoichiometry or from impurities; rather they may result from mechanical imperfection.[14]

DETAILED ANALYSIS OF PRIMARY PHOTOCURRENT

As explained earlier, when the resistance across the input of the preamplifier is reduced to the order of 1000 Ω or less, the observed response is closely proportional to current instead of to integrated charge. The displacement of n electrons a distance dx within a crystal of thickness d induces a charge on the electrodes $dq = ne\,dx/d$. This displacement occurs in a time dt, so that the current $i(t) = dq/dt$ and

$$i(t) = \frac{nev}{d} \tag{8}$$

where v is the drift velocity of the electrons. In the case that n_0 electrons and holes are released at time $t = 0$ close to the negative electrode, electrons only are swept into the crystal, whereas the positive charge makes little contribution. Assume a uniform trap distribution for electrons and a characteristic time T before trapping (the probability per sec, for trapping is taken as constant $A = 1/T$). In this simple case

$$i(t) = \frac{n_0 ev}{d} \cdot e^{-t/T}. \tag{9}$$

The first pulse in Fig. 9 shows such an exponential decay of photocurrent in a region of a crystal where T is of the order of 2 μsec (360 mμ light, absorption constant $K \sim 10^3 \text{cm}^{-1}$). The remaining pulses are at successively higher collecting voltage and

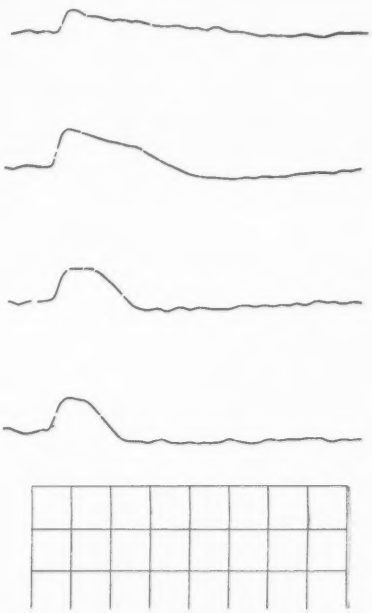


FIG. 9. Current pulses in crystal X_1 annealed in air on MgO. Strongly absorbed 360 mμ light was used with the illuminated electrode negative. This is not a homogeneous crystal, but contains a region adjacent to the surface having a lifetime much longer than that in the interior. First pulse at the top for $V = 10$ V; the second pulse at 21 V shows trapping at about 0·4 mm into the crystal. The third and fourth pulses are at 47 and 60 V, respectively. A graticule having 1 μsec divisions is also shown. Some homogeneous crystals such as I_4 showed exponential decays at higher voltages with much better signal-to-noise ratios. (See Fig. 10).

* Results of spectrographic analysis kindly conveyed to the author by Dr. F. URBACH, Eastman Kodak Laboratories.

show the trapping of electrons at a boundary 0·4 mm into the crystal separating regions of high- and low-electron lifetime. This surface region of long T is the improved layer produced by the air-annealing of a vacuum-grown crystal at 400°C on MgO powder.

A more detailed analysis takes the light-pulse shape and duration into account. Equation (8) shows that the current as a function of time depends upon $n = n(t)$. Direct recombination processes will be unimportant for a low density of electrons swept into the volume of the crystal, and if we assume no shallow trapping states:

$$\frac{dn}{dt} = F(t) - nA(t) \tag{10}$$

$$n(t) = \int_0^t F(\alpha)\,d\alpha - \int_0^t n(\alpha)A(\alpha)\,d\alpha. \tag{11}$$

In these equations $F(t)$ is the number of excitations per sec and $A(t)$ is the probability per sec for trapping. As the basis of our model, let us assume a uniform crystal with constant $A = 1/T$. A solution of equation (11) for a rectangular input function of known duration is given elsewhere.[15] Fig. 5 shows that our excitation or input function can be closely fitted with an equation of the following form:

$$F(t) = F_0 \frac{t^2}{2} e^{-t/T_2}. \tag{12}$$

Substituting this into equation (11) and making use of Laplace transformations, the result may be stated as

$$i(t) = \frac{evF_0}{2d\beta^3}[(\beta^2t^2 - 2\beta t + 2)e^{\beta t} - 2]e^{-t/T} \tag{13}$$

where

$$\beta = \left(\frac{1}{T} - \frac{1}{T_2}\right).$$

Two limiting forms of equation (13) are of interest. First for the case of very short light duration:

$$i(t) = \frac{(eFT_2^3)}{d} ve^{-t/T} \qquad T_2 \ll T \tag{14}$$

which agrees with equation (9), since $eFT_2^3 = n_0 e$.

On the other hand, for short lifetime

$$i(t) = \frac{eF}{2d} \cdot vTt^2 e^{-t/T_2} \qquad T_2 \gg T \tag{15}$$

which has the form of the excitation function.

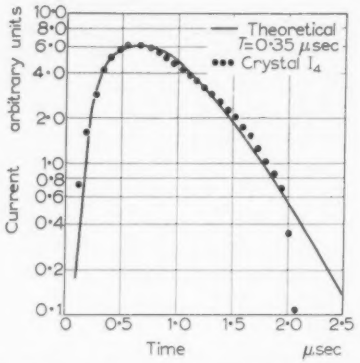


FIG. 10. Current response of a homogeneous crystal no. I_4 at $V = 800$ V for 360 mμ light (nominal) as shown in Fig. 5. The points are taken from a photograph of a single oscillscope trace. The solid line is a normalized plot of equation (13) for $T_2 = 0{\cdot}185$ μsec and $T = 0{\cdot}35$ μsec.

Many crystals have a lifetime comparable with T_2, and Fig. 10 shows a plot of the exact equation (13) normalized to fit the data of crystal I_4 (annealed in vacuum), using a proper choice of the parameter T. The agreement with this simple model is really quite good in as much as the light pulse shape can be fitted analytically. Notice that at a time of the order of 1·88 μsec collection of the electrons at the positive surface is observed. From the applied voltage and thickness of the crystal, electron mobility can be estimated.

Samples cut from boule XVIII annealed in air on quartz powder showed electron lifetimes of the order of 10 μsec. This permits a measurement of drift mobility at low voltage, where the effects of polarization are most certainly negligible. For example, crystal $XVIII_3$ was of thickness $d = 0{\cdot}140 \pm 0{\cdot}002$ cm. At $25{\cdot}5 \pm 2$°C a discontinuity in $i(t)$, corresponding to collection of electrons at the anode, occurred at $T = 8 \pm 0{\cdot}6$ μsec, for an applied voltage $V = 53 \pm 2{\cdot}0$ V. Mobility, μ, is readily calculated from $\mu = d^2/VT$ and turns out to be 46 ± 4 cm²/Vsec. This is in good agreement with the results of HAYNES and SHOCKLEY,[12] who measured

Hall- as well as drift-mobility in the room-temperature range. Since our drift mobility and the values of Hall mobility agree, we conclude that electrons do not spend a very large fraction of the time, T, in shallow trapping states, a result borne out by the fact that current decay in most of our long-lifetime crystals can be represented by a single exponential function.

Electrical-glow-curve experiments on crystals such as $XVIII_3$ indicate that traps of thermal depth about 0·4 eV are present in highest concentration. The lifetime of an electron in a trap of this depth is of the order of 1 msec at 300°K. Therefore, these states control the lifetime of electrons in our fast-pulse measurements as well as in experiments carried out at 77°K. Crystals which have a large range or lifetime at 300°K appear to have similar properties when cooled slowly to liquid-nitrogen temperatures.

THE QUESTION OF HOLE CONDUCTIVITY

When an ultraviolet filter is used, light releases electrons and holes close to the illuminated electrode. By reversing the polarity of the applied voltage, one can arrange to sweep either holes or electrons into the volume of the crystal. Fig. 11 shows the maximum value of current pulses plotted against applied voltage for crystal I_4. It is seen that, under conditions for which electrons would drift into the volume, the current increases with increasing voltage due to increasing drift velocity. Under reverse voltage however (illuminated electrode positive), the response is of opposite sign and shows a saturation behavior. Different samples yielded very similar curves. Fig. 11 may be explained in alternative ways.

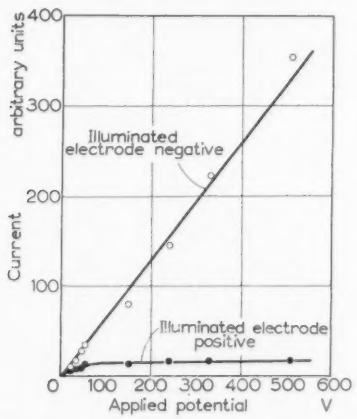


FIG. 11. Current pulse height vs. applied voltage for both polarities of applied voltage. Crystal I_4, strongly absorbed light. With illuminated electrode positive, the response is much smaller, of opposite sign, and shows saturation.

(1) Let us assume that electrons have a long lifetime, but that holes drift a much shorter distance in the crystal and contribute little to the response at either polarity. The saturation behavior of positive polarity response can be fitted quite closely with a theory which assumes exponential absorption of the light and drift of electrons toward the positive illuminated electrode.* For this purpose it is necessary to assume an average light absorption coefficient of 335 cm^{-1}. Although the peak in transmission of the Corning 5860 filter is at 360 mμ, the optical path includes a semiconducting glass electrode and a Mylar film. The combined transmission is shifted toward 365 mμ and includes considerable relative transmission to wavelengths as long as 390 mμ. Furthermore, the spark light source has stronger emission in the range 370–390 mμ than in the range 340–370 mμ. It appears that the observed relative response for opposite polarities is in reasonable agreement with a theory of electron drift only. On this basis there is room for an unsaturated hole current of about 2 units at 500 V. For like yield of electrons and holes, the saturation theory gives an upper limit of hole range in crystal I_4 at 500 V of about $w_h = 0{\cdot}012\ w_e$, corresponding to a hole lifetime of $4{\cdot}2\times10^{-9}$ sec (assuming equal electron and hole mobilities).

(2) On the other hand, evidence has been given for the mobility of holes[17] under certain circumstances. Any explanation of Fig. 11 on the basis of the drift of holes must include the saturation behavior. It would be necessary to evoke surface recombination processes which depend upon collecting voltage or a nonuniform distribution of hole traps in the volume of all samples tested.

* Current pulses for negative polarity in Fig. 11 arise under conditions for which $T \gg T_2$, so that equation (14) applies. On the other hand, the lifetimes for electrons swept toward the illuminated positive electrode satisfy the condition $T \ll T_2$, at least above 50 V, so requiring the use of equation (15). A somewhat better procedure is to observe the charge, q, as a function of V and compare positive and negative polarities with theoretical saturation curves.

Explanation (1) is more satisfactory. Holes have a very much shorter range than electrons in the long-lifetime samples. This result applies to the fast photoresponse of previously unilluminated crystals for low-level pulses of light at room temperature. The possibility that holes have a low effective mobility such as is produced by shallow trapping states is not ruled out. Further experimental work is under way to extend the data to shorter wavelengths and to measure carefully the quantum yield of carriers.

Acknowledgements—The author is grateful for the cooperation of Dr. F. URBACH, his group and Dr. F. HAMM of the Eastman Kodak Laboratories. It is also a pleasure to mention the assistance of MR. O. MCFARLAND and Mr. ROLAND HANSON.

REFERENCES

1. VAN HEERDEN P. J. *Phys. Rev.* **106,** 472 (1957).
2. HECHT K. *Z. Phys.* **77,** 235 (1932).
3. LEHFELDT W. *Nachr. Ges. Wiss. Göttingen* **1,** 171 (1935).
4. BROWN F. C. *Phys. Rev.* **97,** 355 (1955).
5. STÖCKMANN F. *Photoconductivity Conference* (Edited by R. G. BRECKENRIDGE, B. R. RUSSELL and E. E. HAHN). p. 269. Wiley, New York (1956).
6. HILSCH R. and POHL R. W. *Z. Phys.* **108,** 55 (1937); **112,** 252 (1939).
8. MOTT N. F. and GURNEY R. W. *Electronic Processes in Ionic Crystals* p. 122. University Press, Oxford (1948).
9. KOCH E. and WAGNER C. *Z. phys. Chem.* B **38,** 295 (1937).
10. CHASE R. L. and HIGINBOTHAM W. A. *Rev. Sci. Instrum.* **23,** 34 (1952).
11. HEDGES J. M. and MITCHELL J. W. *Phil. Mag.* **44,** 357 (1953).
    CLARK P. V. McD and MITCHELL J. W. *J. Photogr. Sci.* **4,** 1 (1956).
12. HAYNES J. R. and SHOCKLEY W. *Report of a Conference on Strength of Solids* p. 151. Physical Society, London (1948).
    HAYNES J. R. and SHOCKLEY W. L. *Phys. Rev.* **82,** 935 (1951).
13. NAIL N., MOSER F., GODDARD P. E. and URBACH F. *Rev. Sci. Instrum.* **28,** 275 (1957).
14. KANSAKI H. *Phys. Rev.* **99,** 1888 (1955).
15. WARFIELD G. *A Study of Electron-Bombardment-Induced Conductivity in* AgCl *Single Crystals.* Thesis, Cornell University (1950).
16. MOSER F. and URBACH F. *Phys. Rev.* **102,** 15, 19 (1957).
17. LUCKEY G. W. and WEST W. *J. Chem. Phys.* **24,** 879 (1956).
    SMITH L. P. *Semiconducting Materials* (Edited by H. K. HENISCH). p. 114 Butterworths (1951).
    HAMILTON J. F., HAMM F. A. and BRADY L. E. *J. Appl. Phys.* **27,** 874 (1956).
    YAMANAKA C. and SUITA T. *Technol. Rep. Osaka Univ.* **5,** No. 141 47 (1955).

VOL. 4 1958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 217–222.

CATION SUBSTITUTIONS IN $BaFe_{12}O_{19}$

A. H. MONES and E. BANKS

Department of Chemistry, Polytechnic Institute of Brooklyn, N.Y.

(*Received* 29 *July* 1957)

Abstract—The variation in magnetic intensity of $BaFe_{12}O_{19}$ was studied as a function of the substitution of ions such as Al^{III}, Ga^{III}, Cr^{III}, and Zn^{II} for Fe^{III}. All the substituted ions, with the exception of Cr^{III}, were nonmagnetic. The fields employed were not sufficient for magnetic saturation of the polycrystalline preparations; single crystals of some preparations, principally in the Ga^{III} series, were magnetically saturated. From the magnetic saturation and nonsaturation data, some tentative conclusions are drawn regarding the preference of the substituted ions in the available sites, which are octahedrally and tetrahedrally surrounded by oxygen in the structure.

INTRODUCTION

WENT *et al.*[1] deduced a theoretical magnetic saturation value for $BaFe_{12}O_{19}$ at 0°K, based on Néel's hypothesis of ferrimagnetism and Anderson's theory of superexchange. $BaFe_{12}O_{19}$ has the hexagonal structure of the naturally occurring mineral magnetoplumbite,

$$Pb[Fe_{7.5}Mn_{3.5}Al_{0.5}Ti_{0.5}]O_{19}.$$

There are two formula units in the unit cell, which contains two spinel blocks, sandwiched between layers containing barium. In each spinel block there are seven ferric ions in octahedral and two in tetrahedral sites. In and near the barium-containing layers there are two ferric ions per unit cell in trigonal sites and four in octahedral sites. $BaFe_{12}O_{19}$ is described as ferrimagnetic, since the net magnetic moment results from an incomplete cancellation of moments of the ferric ions. The moments from the octahedrally surrounded ferric ions in the spinel blocks and those in the trigonal sites are opposed by a minority of ferric ions in tetrahedral sites in the spinel blocks, along with those in octahedral sites near the barium plane. It would be reasonable to suppose that replacement by nonmagnetic ions of ferric ions in, for example, the tetrahedral sites in the spinel blocks would lead to an increase in magnetic moment of the compound, since the dominance of moments from the octahedral spinel sites and trigonal sites would thereby be enhanced. Similarly, replacement in the latter sites by nonmagnetic cations would reduce the magnetic intensity of the compound.

Al^{III}, Cr^{III}, Ti^{IV}, and Mg^{II} show a preference for octahedral co-ordination in the iron spinels, while Zn^{II} and Ga^{III} prefer tetrahedral co-ordination in these materials.[2]

A related investigation was conducted by VAN UITERT and SWANEKAMP[3] on substitutions of nonmagnetic ions in $BaFe_{12}O_{19}$ and similar compounds. The measurements reported were made at room temperature. Our investigation includes measurement of the low-temperature magnetic properties of the substituted materials and measurements on single crystals.

EXPERIMENTAL PROCEDURE

Mixtures of reagent-grade metal oxides or carbonates were ground in a mortar and pestle in an acetone paste until the mixture appeared homogeneous. The mixtures were fired in platinum crucibles at 1350°C in a flowing oxygen system. The firing time was generally 3½ hr and the samples were quenched in air.

Single crystals of the materials were grown from a sodium carbonate flux. Na_2CO_3 was mixed with the starting materials in the ratio of 2·5 moles of Na_2CO_3 per mole of compound composition. The mixture was introduced into the furnace at 1150°C and reached 1350°C in 2½ hr. At this point cooling was initiated at the rate of 20°C/hr. At 900°C the cooling rate was increased to 60°C/hr for the sake of convenience. Firing was conducted in an oxygen atmosphere. The samples were removed from the furnace at 300°C. The single crystals obtained were washed in hot water and 1/10 HCl to remove the flux and other materials on the crystal surface. The crystals were highly lustrous and shaped in

some cases like perfect hexagons. Single crystals weighing about 20 mg were used for magnetic measurements.

The preparations were investigated by X-ray film and diffractometer techniques.

The preparations containing Ga^{III} were conveniently analyzed with the Philips X-ray fluorescence unit.

The magnetic measurements were made by a modification of the ponderometer technique of RATHENAU and SNOEK.[4] The ponderometer technique involves the measurement of the period of oscillation of a pendulum, containing the magnetic sample, which oscillates in an inhomogeneous symmetrical field. To facilitate this measurement, we introduced an automatic counting system. The pendulum was provided with a short opaque flag, which interrupted a light beam, directed at a cadmium sulfide crystal photocell from a 15 W bulb, once every oscillation. A thyratron circuit, actuated by the photocell, in turn pulsed an electromechanical counter. Spring electrical contacts were placed across the counter dial, a portion of which, corresponding to a fixed number of counts, was made conductive with aluminum tape. The spring electrical contacts, which were simply 24 B. and S. gauge phosphor bronze, completed a circuit through an electric timer when they were across the conductive tape. Thus a fixed number of counts was automatically and dependably timed.

X-RAY RESULTS FOR THE Al^{III} AND Ga^{III} SERIES

The results of the X-ray investigation indicated that a continuous series of solid solutions of the type $BaFe_{12-x}Me_x{}^{III}O_{19}$ (where $Me^{III} = Al^{III}$ or Ga^{III}) formed, at least up to $x = 8$. It is known that $BaAl_{12}O_{19}$, which has the magnetoplumbite structure, can be prepared.[5]

A monotonic decrease of the d_{220} spacing was observed with increasing Ga^{III} concentration. The variation in this reflection was followed on the Philips X-ray diffractometer, using cobalt radiation. The same reflection for the Al^{III} series was broad and too poorly defined for its position to be accurately determined in this manner. Fig. 1 shows the variation of a_0 determined from d_{220} with increasing Ga^{III} content. The decrease is consistent with substitutional solid solution and the fact that Ga^{III} is a smaller ion than Fe^{III}. It is not clear why the lattice constant does not decrease beyond $x = 7$, since for $x = 8$ the presence of starting materials was not observed, nor was a change in structure apparent from the X-ray investigation.

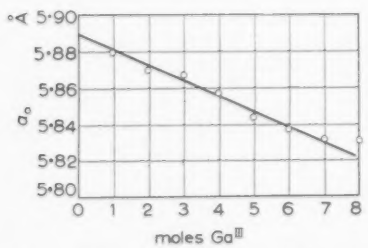


FIG. 1. Variation in a_0 with Ga^{III} content.

MAGNETIC MEASUREMENTS ON THE Al^{III} AND Ga^{III} SERIES

The values obtained from the magnetic measurements on the polycrystalline samples were not saturation values. The measurements were taken from liquid-nitrogen temperature (78°K) to above the Curie temperature in some cases. Fig. 2 shows the variation in Curie temperature with gallium and aluminum substitution. The values for the Curie temperatures are probably high and within 15–20°K of the true value because of the uncertainty in extrapolation of the Curie-temperature plots. It will be noted that the values for the aluminum compounds are higher than for the gallium compounds. This manifests itself sharply at room temperature, where weak magnetic activity is discernible for $BaFe_4Al_8O_{19}$, while $BaFe_4Ga_8O_{19}$ is magnetically inert.

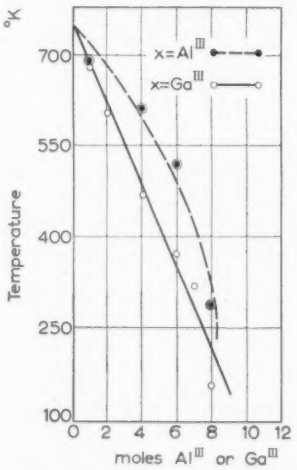


FIG. 2. Curie temperature vs. moles Ga^{III} or Al^{III}.

The thermomagnetization curves obtained for $BaFe_{12}O_{19}$ and the substituted compounds were linear from liquid-nitrogen temperature to about the Curie temperature.

VOL. 4 1958

Table 1 shows the specific magnetization values (in G cm³/g) at liquid-nitrogen temperature for some of the preparations in the series. The magnetic measurements are subject to an error of about 1 per cent.

Table 1. Specific magnetization values for Ga^{III} *and* Al^{III} *substituted solid solutions*

x	$\sigma_{78°K}$ Ga^{III} (G cm³/g)	$\sigma_{78°K}$ Al^{III} (G cm³/g)
0·00 ($BaFe_{12}O_{19}$)	70·3	70·3
0·02	70·4	70·5
0·08	71·3	70·2
0·20	71·0	68·0
0·40	66·8	62·0
0·60	65·8	60·6
1·00	65·7	55·8
2·00	56·2	—
4·00	33·3	28·9

The value $\sigma_{78°K}$ for the compound

$$BaFe_{11}Ga_{0.5}Al_{0.5}O_{19}$$

was 61·7 G cm³/g, which is exactly intermediate between the values obtained for the gallium and aluminum compounds having $x = 1$. Slight increases in σ, as compared to $BaFe_{12}O_{19}$, were observed for some of the solid solutions of very low gallium content, whose values were higher than those of the corresponding Al^{III} preparations.

Table 2. Magnetization values for single crystals in the Ga^{III} *and* Al^{III} *substituted compounds*

Compound	$\sigma_{78°K}$ (G cm³/g)	Room-temperature anisotropy constant (ergs/cm³)
$BaFe_{12}O_{19}$	101·9	$3·3 \times 10^6$
$BaFe_{11·89}Ga_{0·11}O_{19}$	101·7	
$BaFe_{11·78}Ga_{0·22}O_{19}$	101·5	
$BaFe_{11·76}Ga_{0·24}O_{19}$	103·2	
$BaFe_{11·60}Ga_{0·40}O_{19}$	98·6	$3·5 \times 10^6$
$BaFe_{11·54}Ga_{0·46}O_{19}$	93·5	
$BaFe_{11·42}Ga_{0·58}O_{19}$	93·2	
$BaFe_{11·20}Ga_{0·80}O_{19}$	—	$2·9 \times 10^6$
$BaFe_{11·02}Ga_{0·98}O_{19}$	89·3	
$BaFe_{11·8}Al_{0·2}O_{19}$	98·9	$3·3 \times 10^6$
$BaFe_{11·6}Al_{0·4}O_{19}$	81·1	$2·9 \times 10^6$

Table 2 shows the magnetic saturation values for single crystals containing gallium and aluminum at 78°K. The gallium content was determined analytically by X-ray fluorescence analysis. The gallium content in the crystals was different from (and usually less than) that intended in the preparation. Analysis of the crystals containing aluminum was not possible by the above method without considerable modification of the apparatus. Wet analysis for aluminum was not performed, since it is both destructive and very difficult.

The homogeneity of the crystals was not determined. It is anticipated that X-ray work performed on different portions of the crystals will indicate the deviations from homogeneity of composition. However, the similarity in the shapes of the thermomagnetic curves for single crystals and powders of the same composition suggests that the crystals are quite homogeneous. The single crystals could only be magnetically saturated in the direction of easy magnetization, which is parallel to the crystallographic *c*-axis. The materials show strong magnetocrystalline anisotropy. The anisotropy constant, which is a measure of the work required to turn the magnetization from the direction of easy magnetization, was obtained from plots of the magnetization values in the easy and hard (perpendicular to the *c*-axis) direction vs. field strength. The value for $BaFe_{12}O_{19}$ is about 10 per cent higher than that obtained by WENT *et al.*[1] the discrepancy is probably due to inaccuracy in determining the field strength.

The variation in density for the gallium and aluminum compounds was not taken into account in computing their anisotropy constants. It would be expected that the densities of the gallium compounds are higher than that of $BaFe_{12}O_{19}$, since the unit cell volume decreases with gallium concentration, and the unit-cell mass increases. Correcting for increased density would raise the anisotropy constant. The decrease in mass on aluminum substitution is probably accompanied by a decrease in cell volume which would tend to leave the density only slightly affected. At low concentrations of aluminum and gallium, the effect is probably not serious. The change in anisotropy constant from that of $BaFe_{12}O_{19}$, at relatively low concentrations of aluminum and gallium, probably does not significantly affect the magnetization results.

Figs. 3(a)-(c) are typical thermomagnetization

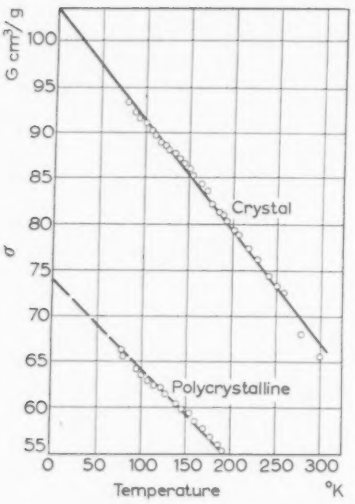


FIG. 3(a). Specific magnetization (σ) vs. temperature curve for $BaFe_{11.42}Ga_{0.58}O_{19}$ crystal and powder.

curves for some of the single-crystal members of this series. Fig. 3 shows the plot for a single crystal and a polycrystalline sample obtained by powdering the crystal. It can be seen that the non-saturation values for the polycrystalline sample are

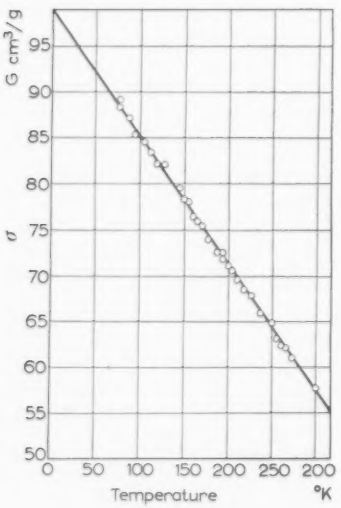


FIG. 3(b). As Fig. 3(a) for $BaFe_{11.02}Ga_{0.98}O_{19}$ crystal.

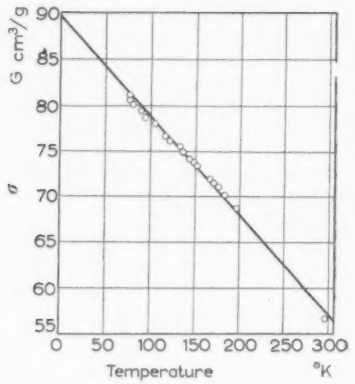


FIG. 3(c). As Fig. 3(a) for $BaFe_{11.6}Al_{0.4}O_{19}$ crystal.

roughly a constant fraction of the saturation values obtained for the crystal.

The saturation value obtained for $BaFe_{12}O_{19}$ at 78°K is in close agreement with the value of WENT *et al.*[1] at this temperature and is very nearly the theoretical value deduced for $BaFe_{12}O_{19}$ at 0°K. The value of σ obtained at 78°K for $BaFe_{12}O_{19}$ is indeed in close agreement with that obtained when thermomagnetization plots down to liquid-hydrogen temperature are extrapolated to 0°K.[2] The curve evidently approaches 0°K with zero slope.

To obtain some idea concerning the location of the gallium and aluminum substitutions, it was assumed that the $\sigma_{78°K}$ values for the substituted compounds could be taken for their saturation values at 0°K. It was also assumed that the $\sigma_{78°K}$ values for the polycrystalline materials were proportional to their saturation values. On this basis, some tentative conclusions could be drawn regarding the nature of the Al^{III} and Ga^{III} substitution in $BaFe_{12}O_{19}$.

The ratio in Table 3 $\dfrac{\text{magnetization substituted}}{\text{magnetization } BaFe_{12}O_{19}}$ compares the values of the magnetization values deduced or observed for the substituted compound with the theoretical value for $BaFe_{12}O_{19}$, 40 Bohr magnetons per unit cell. The column marked "preferential" is for gallium in tetrahedral or aluminum in octahedral spinel sites. Random substitution is for 2/9 of the substituent in tetrahedral and 7/9 in octahedral or fivefold sites.

Table 3. Analysis of the magnetic data

Compound	Magnetization Substituted / Magnetization $BaFe_{12}O_{19}$		
	Preferential	Random	Observed
$BaFe_{12}O_{19}$	—	—	1·00
$BaFe_{11\cdot6}Ga_{0\cdot4}O_{19}$	1·10	0·95	0·94
$BaFe_{11\cdot6}Al_{0\cdot4}O_{19}$	0·90	0·95	0·87
$BaFe_{11}GaO_{19}$	1·25	0·86	0·93
$BaFe_{11}AlO_{19}$	0·75	0·86	0·76
$BaFe_{10}Ga_2O_{19}$	1·50	0·59	0·81
$BaFe_8Ga_4O_{19}$	—	0·33	0·50
$BaFe_8Ga_4O_{19}$	0·00	0·33	0·37
*$BaFe_{11\cdot89}Ga_{0\cdot11}O_{19}$	1·03	0·98	1·00
*$BaFe_{11\cdot78}Ga_{0\cdot22}O_{19}$	1·06	0·97	1·02
*$BaFe_{11\cdot74}Ga_{0\cdot26}O_{19}$	1·07	0·96	1·04
*$BaFe_{11\cdot60}Ga_{0\cdot40}O_{19}$	1·10	0·95	0·95
*$BaFe_{11\cdot54}Ga_{0\cdot40}O_{19}$	1·11	0·94	0·93
*$BaFe_{11\cdot42}Ga_{0\cdot58}O_{19}$	1·15	0·92	0·92
*$BaFe_{11\cdot02}Ga_{0\cdot98}O_{19}$	1·25	0·87	0·89

* Single crystals.

The comparison shows that the observed values were generally consistent with aluminum substituting preferentially in octahedral spinel sites and gallium at low concentrations in tetrahedral spinel sites.

For true saturation values obtained from the crystals, the above comparison shows that, beyond concentrations of about 0·3 g atom gallium, the behavior of the samples was consistent with random substitution in the spinel blocks. Below 0·3 mole most of the gallium appears to enter the tetrahedral positions in the spinel blocks.

MAGNETIC MEASUREMENTS WITH THE SUBSTITUTION OF Zn^{II}, Mg^{II}, Ti^{IV}, AND Cr^{III}

Less extensive substitutions were made in $BaFe_{12}O_{19}$ with Zn^{II}, Mg^{III}, Ti^{IV}, and Cr^{III}. The zinc ion shows a very strong preference for tetrahedral co-ordination in the spinels and other compounds. The preparations, along with their $\sigma_{78°K}$ values, are listed in Table 4. The formulas merely represent the starting compositions; no implication as to the composition of the final product is intended.

These values were obtained for polycrystalline materials and are not saturation values. Making the same assumptions as above with regard to the

Table 4. Magnetic measurements on other substituted compounds

Compound	$\sigma_{78°K}$
$BaFe_{11\cdot6}Zn_{0\cdot6}O_{19}$	78·5
$BaFe_{11}ZnO_{18\cdot5}$	82·1
$BaFe_{11}Zn_{1\cdot5}O_{19}$	80·3
$BaFe_{11\cdot6}Mg_{0\cdot4}O_{19}$	65·4
$BaFe_{11}TiO_{19\cdot5}$	58·0
$BaFe_{10}Cr_2O_{19}$	40·2

$\sigma_{78°K}$ values, it was found that the effect of zinc substitution was consistent with tetrahedral preference, while the other ions behaved as if they preferred octahedral sites in the spinel blocks. The magnetization values for the zinc compounds are significantly higher than for $BaFe_{12}O_{19}$ or any of the substituted compounds. Of the zinc compounds, only $BaFe_{11}ZnO_{18\cdot5}$ appeared to have the pure $BaFe_{12}O_{19}$ structure $BaFe_{11\cdot6}Zn_{0\cdot6}O_{19}$ showed extra X-ray reflections and $BaFe_{11}Zn_{1\cdot5}O_{19}$ showed some X-ray reflections of zinc oxide. The compounds containing zinc were cooled slowly (100°C/hr) from temperature instead of being quenched, to inhibit the "freezing in" of zinc in octahedral sites.[6] Similar slow cooling of gallium-substituted compounds did not affect their magnetic intensity.

The recent work of JONKER *et al.*[7] showed compounds corresponding to $Ba_2Me_2Fe_{12}O_{22}$, $Ba_3Me_2Fe_{24}O_{41}$, and $BaMe_2Fe_{16}O_{27}$ (where Me represents zinc and magnesium, among other divalent metal ions). It is possible that some of the unidentified lines in the X-ray patterns of samples containing zinc arise from the presence of one or more of these materials.

Compounds of the type

$$Ba_{1-x}Me_x^{II}La_x^{III}Fe_{12-x}^{III}O_{19}$$

were prepared, where $Me^{II} = Zn^{II}$ or Mg^{II}. The lanthanum was introduced to compensate for the decrease in positive charge occasioned by the introduction of divalent ions in place of ferric ions. Measurements showed that when $Me^{II} = Mg^{II}$ a decrease occurred. Additional preparations of this type are presently in progress, where Me^{II} represents ions other than Mg^{II} or Zn^{II}. It has been shown from other work that lanthanum substitutes for Ba^{II} in $BaFe_{12}O_{19}$.[8]

CONCLUSIONS

From magnetic saturation and nonsaturation data on substituted $BaFe_{12}O_{19}$, it is tentatively concluded that gallium tends primarily to substitute at random in the available sites in the spinel blocks and aluminum in the octahedral sites in the spinel blocks. Zinc, substituting for Fe^{III}, increases the magnetic intensity of $BaFe_{12}O_{19}$ in accord with preferential tetrahedral substitution in the spinel blocks, while Ti^{IV}, Mg^{II}, and Cr^{III} decrease the magnetization, in accord with octahedral preference in the spinel blocks.

REFERENCES

1. Went J. J., Rathenau G. W., Gorter E. W. and van Oosterhout G. W. *Philips Tech. Rev.* **13,** 194 (1952).
2. Gorter E. W. Thesis, University of Leyden (1954).
3. Van Uitert L. G. and Swanekamp F. W. *J. Appl. Phys.* **28,** 482 (1957).
4. Rathenau G. W. and Snoek J. L. *Philips Res. Rep.* **1,** 239 (1946).
5. Adelskold V. *Ark. Kemi Min. Geol.* **12A,** No. 29, 1 (1938).
6. Brockman F. G. *Phys. Rev.* **77,** 841 (1950).
7. Jonker G. H., Wijn H. P. S. and Braun P. B. *Philips Tech. Rev.* **18,** 145 (1956).
8. Summergrad R. N. and Banks E. *J. Phys. Chem. Solids* **2,** 312 (1957).

VOL.
4
1958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 223–226.

INELASTIC MAGNETIC SCATTERING OF NEUTRONS AT HIGH TEMPERATURES

P. G. DE GENNES

Centre d'études nucléaires de Saclay, Gif-sur-Yvette (S. et O.), France

(*Received* 29 *June* 1957)

Abstract—The magnetic inelastic scattering of neutrons by paramagnetic media with exchange interactions has been studied, using the time-dependent spin correlation functions introduced by VAN HOVE. It is found that the even moments of the energy transfer depend strongly on momentum transfer (for small momentum transfer, all moments go to zero), and on the state of the scatterer (single crystal or powder).

The shape and half-width of the spectrum are given as functions of the scattering angle for a polycrystal with nearest-neighbour interactions. The whole study has been confined to the high-temperature region.

METHOD OF MOMENTS

WE are dealing here with paramagnetic insulators, in which the exchange coupling is assumed to be described by the Hamiltonian

$$\mathscr{H} = \sum_{RR'} J_{RR'} \boldsymbol{S}_R \cdot \boldsymbol{S}_{R'}. \tag{1.1}$$

As first noted by VAN VLECK,[1] for temperatures much higher than the paramagnetic Curie point Δ, such an interaction causes an almost random motion of each spin in the crystal, and this makes the magnetic scattering of slow neutrons dominantly *inelastic*.

The 1938 verifications of that fact were only through integrated cross-sections. However, experiments are now being carried out at Chalk River[2] and Saclay[3] which give some information on the energy spectrum of the scattered neutrons. It is therefore useful to make a theoretical study of these effects.*

For a rigid crystal above the Curie point, the magnetic neutron cross-section can be written, in VAN HOVE's notation[4]:

$$\frac{d\sigma}{d\Omega d\omega} = N\left(\frac{e^2}{mc^2}1{\cdot}91\right)^2 \frac{k_1}{k_0}|F(\boldsymbol{\varkappa})|^2 \frac{1}{3\pi} \times$$

$$\times \int dt \left(\sum_R \exp[i(\boldsymbol{\varkappa}\cdot\boldsymbol{R} - \omega t)] \langle \boldsymbol{S}_0(0)\cdot\boldsymbol{S}_R(t)\rangle\right) \tag{1.2}$$

$\hbar\kappa$ and $\hbar\omega$ are the momentum and energy transfers from neutron to crystal.

In this section, we shall relate this formula to the method of moments introduced by VAN VLECK,[1] using the following symbols:

$$A(\boldsymbol{k}_0, \boldsymbol{k}_1) = \left(1{\cdot}91\,\frac{e^2}{mc^2}\right)^2 \frac{1}{3\pi}\cdot\frac{k_1}{k_0}|F(\boldsymbol{\varkappa})|^2 S(S+1) \tag{1.3}$$

$$U_R(t) = \langle \boldsymbol{S}_0(0)\cdot\boldsymbol{S}_R(t)\rangle [S(S+1)]^{-1} \tag{1.4}$$

$$p_\kappa(\omega) = \sum_R \int dt \cdot \exp[i(\boldsymbol{\varkappa}\cdot\boldsymbol{R} - \omega t)] U_R(t). \tag{1.5}$$

The A factor is a known function of the scattering parameters. The effects of the statistical behaviour of the spin assembly are contained in $p_\kappa(\omega)$. Experimental results are best presented (when possible) in terms of $p_\kappa(\omega)$.

Let us now recall some general properties of the correlation functions $U_R(t)$ which will be useful when studying $p_\kappa(\omega)$. First, in the high-temperature region, correlations between different spins taken at different times are 0:

$$U_R(0) = \delta_{R,0}. \tag{1.6}$$

* A *résumé* of the results reported here has been given in *C.R. Acad. Sci., Paris* **244**, 752 (1957).

Second, at all temperatures, the total spin of the crystal is a constant of the motion, and this implies:

$$\sum_R U_R(t) = \sum_R U_R(0). \tag{1.7}$$

We shall now set up an approximate function $p_\kappa(\omega)$ by looking at the moments:

$$\langle \omega^{2n} \rangle_\kappa = \frac{\int p_\kappa(\omega)\omega^{2n}\, d\omega}{\int p_\kappa(\omega)\, d\omega}. \tag{1.8}$$

Moments of odd order are all 0 in the high-temperature limit. From equation (1.5) the general formula for the even moments is:

$$\langle \omega^{2n} \rangle_\kappa = (-)^n \sum_R \exp(i\boldsymbol{\varkappa}\cdot\boldsymbol{R}) \left[\frac{d^{2n}}{dt^{2n}} U_R(t)\right]_{t=0}. \tag{1.9}$$

The VAN VLECK values[1] for the moments are obtained if it is assumed (arbitrarily) that:

$$U_R(t) \equiv 0$$

for $R \neq 0$. This oversimplification in the VAN VLECK paper can be traced to his equation (2), where a single atomic spin is used in the matrix element of the neutron-lattice interaction. Unfortunately, the moments computed in this way are quite incorrect for small-angle scattering, which was the proposed experimental situation.

We shall see that in equation (1.9), if the range of the exchange force is finite, the sum $\sum_R$ for a given n has but a finite number of terms: each moment is a continuous function of the momentum transfer. Also, from equation (1.7):

$$\langle \omega^{2n} \rangle_{\kappa=0} = 0 \tag{1.10}$$

All moments behave like κ^2 near $\kappa = 0$.

NUMERICAL COMPUTATION OF MOMENTS

To make use of the general formula (1.9) for the moments, one has to compute the derivatives $[d^{2n}/dt^{2n}\, U_R(t)]_{t=0}$. This is done by repeated use of the equations of motion:

$$\hbar \frac{d}{dt}\boldsymbol{S}_R = 2\sum_{R'} J_{RR'}(\boldsymbol{S}_R \wedge \boldsymbol{S}_{R'}). \tag{2.1}$$

After the required number of derivatives have been taken, the averaging process is readily carried out in the high-temperature limit, by application of the following rules:

$$\left.\begin{aligned} &\langle S_R^x(0) S_{R'}^x(0)\rangle = \tfrac{1}{3}\delta_{RR'} S(S+1) \\ &\langle S_R^x(0) S_{R'}^y(0)\rangle = 0 \\ &\langle S_R^x(0) S_{R'}^y(0) S_{R''}^z(0)\rangle \\ &\qquad = \frac{i}{6}\delta_{RR'}\delta_{R'R''} S(S+1). \end{aligned}\right\} \tag{2.2}$$

We shall now restrict ourselves to *nearest-neighbour* interactions, the distance between neighbours being b, and the number of nearest neighbours of a given ion being v.

From equations (2.1) and (2.2) it is readily seen that $(d^2/dt^2\, U_R)_0$ is 0, except when atoms (R) and (O) are identical, or nearest neighbours. Then:

$$[d^2/dt^2\, U_0]_{t=0} = -v[d^2/dt^2\, U_1]_{t=0} = -\tfrac{2}{3}vS(S+1)(2J/\hbar)^2.$$

For a *polycrystal*, after averaging equation (1.9) over angles, we get:

$$\langle \omega^2 \rangle = \tfrac{2}{3} vS(S+1)\left(\frac{2J}{\hbar}\right)^2 \left(1 - \frac{\sin(\kappa b)}{\kappa b}\right). \tag{2.3}$$

We have computed the fourth moment, following the same lines, but only for a *simple cubic* lattice of magnetic ions. For a polycrystal, after averaging over angles, we find:

$$\langle \omega^4 \rangle_\kappa = [S(S+1)]^2 \left(\frac{2J}{\hbar}\right)^4 \left\{ \left(\frac{224}{3} + \frac{2}{S(S+1)}\right) \times \left(1 - \frac{\sin(\kappa b)}{\kappa b}\right) - 16\left(1 - \frac{\sin(\kappa b\sqrt{2})}{\kappa b\sqrt{2}}\right) - 4\left(1 - \frac{\sin(2\kappa b)}{2\kappa b}\right)\right\}. \tag{2.4}$$

Moments of higher orders and related problems are now being studied by HERPIN, LAFORE and LATTÈS.

We now compare $p_\kappa(\omega)$ with a Gaussian distribution. Accordingly, we consider the ratio:

$$q(\kappa) = \frac{3(\langle \omega^2 \rangle_\kappa)^2}{\langle \omega^4 \rangle_\kappa}.$$

VOL. 4 1958

This is unity for a Gaussian. The actual values are plotted against κb in Fig. 1. It can be seen that for $\kappa b \gtrsim \pi$, $q(\kappa)$ is not sensibly different from 1. As

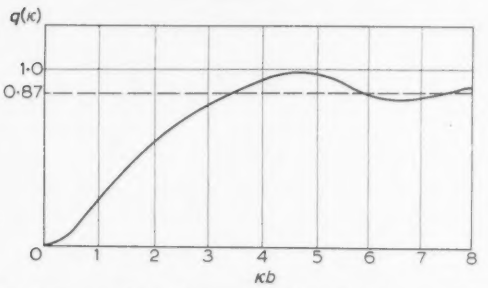


FIG. 1. Comparison of second and fourth moments of the distribution function (for a polycrystal).

we know that all moments of a higher order are finite, it is certainly a sufficient approximation to put:

$$p_\kappa(\omega) = \left[\frac{2\pi}{\langle\omega^2\rangle_\kappa}\right]^{\frac{1}{2}} \exp\left(\frac{-\omega^2}{2\langle\omega^2\rangle_\kappa}\right).$$

On the other hand, for $\kappa b \ll \pi$, $q(\kappa)$ goes to 0; the function is certainly not a Gaussian. This will be checked in the next section.

SMALL MOMENTUM TRANSFERS

We now study the region where $\kappa b \to 0$. The moments (2.3) and (2.4) become:

$$\langle\omega^2\rangle = \frac{v}{9}S(S+1)\left(\frac{2J}{\hbar}\right)^2(b\kappa)^2 \qquad (3.1)$$

$$\langle\omega^4\rangle \cong \frac{27}{6}S^2(S+1)^2\left(\frac{2J}{\hbar}\right)^4(b\kappa)^2. \qquad (3.2)$$

The function $p_\kappa(\omega)$ is best described by a "truncated Lorentz" curve:

$$\left.\begin{aligned} p_\kappa(\omega) &= \frac{2\Lambda_1\kappa^2}{\omega^2+(\Lambda_1\kappa^2)^2} \qquad |\omega| < \omega_c \\ p_\kappa(\omega) &= 0. \qquad\qquad\quad |\omega| > \omega_c \end{aligned}\right\} \qquad (3.3)$$

The values of Λ_1 and ω_c can be deduced from (3.1) and (3.2)

$$\Lambda_1 \cong 0{\cdot}23b^2[S(S+1)]^{\frac{1}{2}}\left(\frac{2J}{\hbar}\right) \qquad (3.4)$$

$$\omega_c \cong 4{\cdot}5[S(S+1)]^{\frac{1}{2}}\left(\frac{2J}{\hbar}\right). \qquad (3.5)$$

The half-width is of the order of:

$$\Lambda_1\kappa^2 \cong \omega_c\left(\frac{\kappa b}{4{\cdot}5}\right)^2 \qquad (3.6)$$

so the cut off ω_c is far away in the wings, and our approximation is justified.

CONCLUSIONS

Half width Γ_κ of the function $p_\kappa(\omega)$. It is plotted against κb on Fig. 2. The low κ region of the curve is valid only for a simple cubic lattice of magnetic

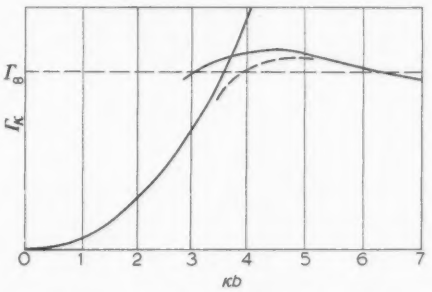


FIG. 2. Half-width of the distribution function as a function of momentum transfer (lattice is assumed simple cubic).

ions. The dotted line results from an approximate interpolation between the low κ and high κ regions.

The high κ limit is:

$$\hbar\Gamma_\infty \cong 2{\cdot}9\frac{k_B\Delta}{[vS(S+1)]^{\frac{1}{2}}}. \qquad (4.1)$$

The small fluctuations of Γ_κ around Γ_∞ in the high κ region are not very significant, because of errors implied by the Gaussian assumption. But the decrease of Γ_κ for $\kappa b < \pi$ is meaningful and might possibly be observed. In the first experiments by BROCKHOUSE[2] this region was not studied, and formula (4.1) is valid (within molecular field approximation for estimation of J).

APPENDIX

Large-angle scattering at all temperatures

We retain the condition $\kappa b > \pi$ (for a polycrystal), so

that the magnetic ions can still approximately be taken to scatter independently.* We also assume that, below the Curie point, magnetization is oriented at random in different domains. Using the same methods as before, we get for the total magnetic cross-section

$$\frac{d\sigma}{d\Omega} = \left(\frac{e^2}{mc^2}1{\cdot}91\right)^2 \tfrac{2}{3}[S(S+1)-(\langle \boldsymbol{S}\rangle_T)^2]. \quad \text{(A.1)}$$

(In this formula, the A factor in the cross-section has been taken as a constant. This can be corrected in an actual case.) The temperature-dependence of equation (A.1) is plotted in Fig. 3. The most interesting region is that immediately under the Curie point, where equation (A.1) varies rapidly.

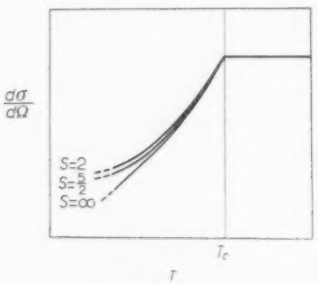


FIG. 3. Temperature-dependence of the incoherent cross-section (after integration over final energies of the neutron). For the approximations involved, see the Appendix.

* Short-range order effects have been discussed elsewhere[5] and shown to be small if $\kappa b > \pi$. It is also assumed that the measurement is *not* carried out in the vicinity of a Bragg peak, to avoid a spurious elastic scattering.

At finite temperatures, $p_\kappa(\omega)$ is no longer symmetrical around $\omega = 0$. The first moment can be shown to be:

$$\hbar\langle\omega\rangle_T = -\frac{E(T)}{S(S+1)-(\langle \boldsymbol{S}\rangle)^2} \quad \text{(A.2)}$$

$$E(T) = \sum_R 2J_{OR}\langle \boldsymbol{S}_O \boldsymbol{S}_R\rangle_T$$

$\frac{1}{2}E(T)$ is the average exchange energy of one of the magnetic ions at temperature T. The origin is such that $E(\infty) = 0$.

We see that the first moment can be connected with the spontaneous magnetization and magnetic specific heat of the crystal. It is small above the Curie point, varies steeply near the Curie point, and, at low temperatures, it reaches the limit:

$$\hbar\langle\omega\rangle_{T=0} \cong \frac{3k_B T_c}{S+1}$$

valid for a ferro- or an anti-ferromagnet with only nearest-neighbour interaction, and within molecular field approximation.

REFERENCES

1. VAN VLECK J. H. *Phys. Rev.* **55,** 924 (1939).
2. BROCKHOUSE B. N. *Phys. Rev.* **99,** 601 (1955).
3. GALULA M. Unpublished work.
4. VAN HOVE L. *Phys. Rev.* **95,** 1374 (1954).
5. DE GENNES P. G. and FRIEDEL J. *J. Phys. Chem. Solids* **4,** 71 (1958).

VOL. 4 1958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 227–238.

CRYSTAL DISTRIBUTION IN VARIOUS POLYETHYLENES

A. CHARLESBY and L. CALLAGHAN

Tube Investment Research Laboratories, Hinxton, Cambs.*

(*Received* 12 *July* 1957)

Abstract—The melting characteristics of various types of polyethylene have been deduced from volume measurements over a range of temperatures. These can be related to a distribution of crystallites of normal paraffins whose melting characteristics reproduce the observed results. This equivalent distribution offers a simple representation of the differences between high- and low-density polyethylenes, either alone or when mixed.

OL.
4
958

INTRODUCTION

THE physical properties of long-chain polymers are often determined by the presence within them of crystalline regions. For such regions to form, a certain degree of regularity must be present both in the main and in the side chain. This is notably the case for nylon and for polytetrafluorethylene. In other polymers, where bulky side-groups are located in an irregular manner along the main chain, crystallization is inhibited, and a glass-like amorphous structure is obtained even at low temperatures (e.g., polystyrene, polyvinyl acetate). Techniques of polymerization have recently been discovered which favour a regular arrangement of side groups on either side of the main chain. Such *isotactic* polymers are highly crystalline, and their properties are sufficiently different from those of the corresponding polymer prepared by the older techniques to warrant their being considered as essentially new types of polymer. The number and size distribution of crystallites has a considerable effect on the mechanical properties of a partly crystalline polymer.

In the case of polyethylene, where no side groups (other than hydrogen) are present in the monomer, crystallinity is partly limited by the occurrence of branched chains introduced by a chain-transfer process during polymerization. These branches may consist of methyl, ethyl, or longer groups, and in the past their evaluation has met with some difficulties. Infra-red methods may be used to determine the number of methyl groups, while exposure to high-energy radiation results in the preferential removal of these side chains (HARLEN *et al.*[1]) which can then be analysed by mass spectroscopy. Other factors likely to influence crystallite sizes include the finite molecular weight of the molecule and the presence of unsaturation, either in the main chain, at the ends of the molecule, or as pendant methylene groups. With the exception of *cis* unsaturation in the main chain, these can be studied by infra-red analysis.

In polyethylene the degree of branching (and hence of crystallinity) is primarily determined by the conditions of polymerization. With the earlier technique involving conditions of high pressure and temperature, the degree of branching is generally higher than in the more recent catalytic techniques introduced by Ziegler and Phillips. Data on some typical specimens are shown in Table 1, in which the values are compared on a common basis of 1000 C atoms. To these figures one should add the number and weight average molecular weights to take into account end-group corrections, but the available data are not sufficiently consistent for this purpose. These factors will affect the size and number of the crystallites present at any temperature as well as their thermal stability. The higher the number of irregularities in the chain, the smaller are the crystallites and the lower the density. The greatest regularity in structure may be expected to occur in the case of polymethylene, obtained by the polymerization of diazomethane, although in practice there does appear to be some branching present in the specimen we have examined. Mention should also be

* Present address: Royal Military College of Science, Shrivenham, Wilts.

made of Alkathene H.D., a polyethylene made by the high-pressure process, with properties intermediate between those of the more highly branched Alkathene and the Ziegler or Marlex material.

Branched polyethylenes can also be prepared under accurately controlled conditions by copolymerizing ethylene with vinyl monomers such as propylene, butene, etc. By varying the relative amounts of the constituents, any required degree of branching of methyl, ethyl, etc., groups can be introduced into the polymer chain, and their effect on the physical properties studied.

A considerable amount of data has been published on the variation of the degree of crystallinity with temperature, most of these data being devoted to the low-density (high-pressure) polyethylene which was available at the time. The methods used included measurement of X-ray diffraction intensity,[3] infra-red spectra,[1] heat content,[5] and density.[6] The results obtained by various authors agree only qualitatively in indicating a degree of crystallinity of the order of 50 per cent at 20°C, decreasing to zero at about 115°C. Quantitative agreement between the various results is, however, poor owing partly to the intrinsic errors of the methods adopted or assumptions made, but also in part to differences in the molecular weight and degree of branching of the various grades of polyethylene samples studied.

X-ray or electron-diffraction measurements[7] of polyethylene show a sharp ring pattern, due to the crystalline regions, and a diffuse halo pattern, arising from the amorphous regions. As the temperature is raised, the crystalline pattern becomes less intense, and the crystal spacings alter very slightly towards a hexagonal structure, though melting intervenes before this is reached. Of interest in the present context is the observation that there is not a slow transition in the pattern, whereby the crystalline pattern becomes more diffuse and spacings tend to those of the amorphous regions; at any stage of the melting process the crystalline and amorphous patterns remain quite distinct. This observation indicates that melting of each crystallite

Table 1. *Unsaturation and branching in three typical polyethylenes**

Process	Density	Type	Unsaturation: C = C per 1000 C					Methyl groups CH_3 per 1000 C	
			Total unsaturation	Main chain trans (a)	Terminal (b)	Side chain pendent (c)	Total methyl ends	Terminal	Branches methyl + ethyl
I.C.I. High Pressure	Low	DYNH	0·6	0·1 (17%)	0·29 (15%)	0·41 (68%)	21·5	4·6	16·9
Ziegler Low Pressure	High	Super Dylan	0·7	0·3 (43%)	0·22 (32%)	0·18 (25%)	3	~ 2	1
Phillips Low Pressure	High	Marlex 50	1·5	1·41 (94%)	0·015 (1%)	0·075 (5%)	< 1·5	< 1·5	< 1
				—C=C—	—C=C	—C— ‖		—CH_3	—C— | CH_3 —C— | C_2H_5
Infra-red peak absorption (*m*)				10·35	10·1, 11·0	11·25		7·25	(Ethyl) 13·0

* Data based on D. C. SMITH.[2]

can be considered as a sudden process, and that the division of the polymer into distinct crystalline and amorphous regions is a justifiable approximation.

Radiation and crystallinity

The degree of crystallinity of polyethylene at any temperature can also be modified by exposure to high-energy radiation in the form of fast electrons, X- or gamma-rays, or atomic-pile radiation. In this case, the effects produced involve not only changes in crystallinity, but also changes in chemical structure (primarily unsaturation) and cross-linking between molecules. An early study of the changes in the crystallinity of high-pressure polyethylene due to atomic-pile radiation was published several years ago (CHARLESBY and ROSS[8]). In the present paper, a comparison is made of the crystallinity in the various forms of polyethylene which have since become available, while in another paper the effect of different types of high-energy radiation on this crystallinity will be assessed.

EXPERIMENTAL

Preparation of specimens

In order to remove air cavities and volatile additives, such as antioxidants, which have been found to interfere with measurements in the 20–180°C temperature range, the polyethylene specimens (each about 2·0 g) were heated in wide (4·5 cm) glass tubes to 250°C for several hours whilst connected to a mercury diffusion pump. After all signs of bubbling had ceased, the specimens were allowed to cool slowly under vacuum and maintained at 20°C for several hours before density measurements were made.

Measurements at 20°C

Several particles of polythene of each type or radiation dose (approximately $1\times0{\cdot}5\times0{\cdot}5$ cm) were floated in a mixture of ethyl alcohol and water, stirred magnetically, and kept at 20°C in a thermostat. From the concentration of alcohol required to suspend the particles for some 15 min, the density of the mixture was determined.[9]

These measurements emphasized the need to outgas the polymer, as otherwise there was a marked difference in the density of different particles.

Measurements between 20° *and* 180°C

A weighed quantity (about 1·0 g) of small slices of polythene (about 0·5 mm thick), previously calibrated at 20°C, was placed in a dilatometer (Fig. 1). After sealing off the side arm, the vessel

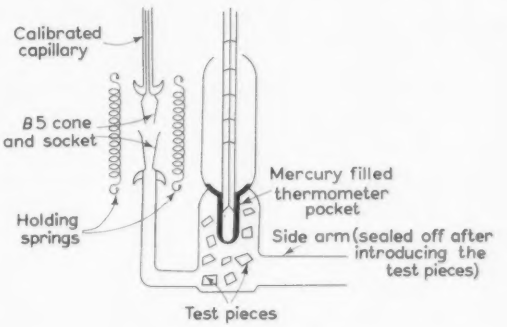


FIG. 1. Dilatometer.

was weighed and filled with redistilled mercury under a mercury-diffusion-pump vacuum (about 10^{-5} mm Hg) to exclude air pockets. The filled vessel was weighed, fitted with a precision-bore capillary, and suspended in an oil bath, which was maintained at 20°C for several minutes before the height of the mercury column was measured with a travelling telescope. The temperature was increased gradually (about 0·5°C/min) and the height of the mercury column measured at suitable intervals. Density calculations were based on the value obtained at 20°C by the flotation method, corrections being made for the expansion of the dilatometer bulb (α Pyrex glass $= 0{\cdot}036\times10^{-4}$ cm^3/cm^3 per °C) and of the mercury.[10]

The reproducibility of the measurements was checked (*a*) by measuring the expansion of mercury only, and comparing the changes in volume with those given in reference tables, errors being of the order of two parts in 10^4 or less, (*b*) by using different rates of heating; even at very high rates (2°/min) the differences in volume were negligible, and (*c*) by repeated heating or cooling, when slight differences were observed only very close to the melting point owing to quenching or annealing. The values given in this paper refer to very slow rates of temperature change, when the material can be considered as substantially annealed. The thin samples surrounded by mercury may be

considered as in good thermal equilibrium, the temperature difference between the centre of each polyethylene particle and its surface being less than 0·2°C.

Observed values of specific volume (1/density) of two samples of Phillips Marlex 50 are shown in Table 2. For any one specimen, total errors are below one part in a thousand.

Table 2. Effect of sample size on specific volume measurement

Temperature (°C)	Specific volume	
	Sample 1*	Sample 2†
20	1·036	1·035
80	1·063	1·063
120	1·100	1·099
160	1·285	1·287

* Small block; † 2 mil film.

DENSITY OF CRYSTALLINE AND AMORPHOUS REGIONS

The structure of polyethylene at room temperature, as revealed by X-ray or electron diffraction, shows the simultaneous occurrence of crystalline and amorphous regions; the former are evidenced by a sharp ring pattern identical with that obtained from normal paraffins in the solid state, the latter by amorphous halos with spacings similar to those obtained from liquid paraffins. It is generally assumed that the whole structure of polyethylene can be represented to a sufficient degree of accuracy by a series of crystals of varying sizes separated by amorphous regions. No accurate information is available on the size of these crystals, but they are believed to be very much smaller than the average length of a polymer molecule. Each molecule therefore traverses a number of such crystalline and amorphous regions and serves to link adjacent crystals together.

The specific volume ($v = 1/\text{density}$) of polyethylene can then be represented by the expression

$$v = fv_c + (1-f)v_a$$

where f is the weight fraction of crystalline material v_c and v_a are respectively the specific volumes of the crystalline regions and of the amorphous regions. Knowing v, v_c, and v_a at various temperatures, it is possible to calculate f as a function of temperature or of chemical structure.

Determination of v_c

Early values by BUNN[11] gave the structure of long-chain paraffin molecules as being orthorhombic with cell dimensions 7·40, 4·93, and 2·534 Å. With four carbon atoms per unit cell, this leads to a specific volume v_a of $0{\cdot}99_3$. Electron-diffraction measurements by CHARLESBY[7] on thin polyethylene films led to a specific volume v_a of $0{\cdot}99_7$. Owing to the thermal agitation, this volume increases with temperature[12] in a similar manner to that of long-chain n paraffins. Our own X-ray measurements on various high- and low-pressure polyethylenes prepared under a variety of conditions (moulded, cast in solution, stretched, etc.) give slightly different spacings for the unit cell, but the volume per four carbon atoms equals 94·1 $\pm$0·2 Å^3 for all except low-molecular-weight polyethylene (Alkathene 200), corresponding to a specific volume (v_c) of 1·01$\pm$0·002, at 20°C. WALTER and REDING[13] have recently reported a variation of unit cell dimensions with degree of branching, which is larger than that given by MATTHEWS *et al.*[14] The values of WALTER and REDING correspond to v_c values of $0{\cdot}99_5$ for Marlex and about $1{\cdot}01_5$ for Alkathene. In the calculations given below, the value assumed throughout for v_c is 1·0146 for Alkathene 7 and Alkathene H.D, 1·0114 for Hostalen F, 1·0103 for Marlex 50, all at 20°C (based on X-ray measurements), with a common thermal expansion coefficient of the crystalline regions of 4×10^{-4}. Others working in this field have chosen slightly different values of v_c (Table 3), resulting in greater differences in the degree of crystallinity. Variations in the position of the X-ray diffraction peak would be of particular importance in X-ray determinations of crystallinity, but are of lesser importance in density determinations.

Determination of v_a

No direct method is available for estimating the specific volume of the amorphous regions in a partially crystalline polyethylene. In this work, two methods have been adopted, in both of which it is assumed that there is no fundamental difference in structure between the amorphous regions

Table 3(a). *Comparison of values of* v_a *and* v_c

v_c	Temp. (°C)	Coefficient of expansion $\times 10^{-4}$	v_c at 20°C	Method	Reference
			1·008		BUECHE[19]
		4	1·010 (Marlex) 1·011 (Ziegler) 1·014 (Alk. 7)	X-ray	This paper
1·025	0	3	1·031		NIELSEN
0·998	25				BRYANT[18]
				X-ray	PRICE
0·993	—			X-ray	BUNN[11]
0·986 0·995 1·001 1·015	25	4	0·985 0·994 1·000 1·014	X-ray	WALTER and REDING[13]
1·00	—			X-ray	MATTHEWS *et al.*[14]
1·00	—	6			HUNTER and OAKES[6]

Table 3(b). *Specific volume of the amorphous regions* v_a

v_a	Temp. (°C)	Coefficient of expansion $\times 10^{-4}$	v_a at 20°C	Method	Reference
1·152	0	7·8	1·1676		BUECHE[19]
1·163	24		1·160	X-ray	KRIMM[3]
		9·59 (8·7–10·2)	1·156 (1·149–1·169)	density	This paper
			1·152	calculated	This paper
1·152	0	6·75	1·166	density	NIELSEN
1·175	—			density	REDING and LOVELL[17]
1·242–1·212	25		1·237–1·207	density	BRYANT[18]
		8·6;9·5	1·169;1·158	density	HUNTER and OAKES[6]
1·282				X-ray	MATTHEWS *et al.*[14]

in partly crystalline polyethylene, polyethylene in the liquid state and liquid paraffins. This assumption is justified by the similarity in their X-ray diffraction patterns and by agreement obtained between the following two independent methods for evaluating v_a.

(a) Above the melting point, polyethylene expands approximately uniformly with temperature. The curve may be extrapolated to derive a value for v_a at any lower temperature.

(b) At any given temperature, the specific volume of liquid *n* paraffins is extrapolated to an infinite molecular weight. Details of this method are given in the Appendix.

The extent of agreement obtained by these two methods is very satisfactory (Table 4), and serves to justify the assumption that v_a and v_c are additive. Some slight differences would be expected to arise from the presence of branches in liquid polyethylene which are not present in shorter linear paraffins, but the observed differences lie well within experimental error (see Appendix). It may be concluded that within experimental error the expansion of the amorphous regions is linear with temperature.

Table 4. Calculated specific volume of amorphous regions (low-density polyethylene)

Temperature (°C)	From liquid *n*-paraffins		From amorphous Polyethylene
	Calculated v_2 (Å³/cm³)	$1/\rho$ (cm³/g)	$1/\rho$ (cm³/g)
20	26·81	1·152	1·156*
100	28·6	1·229	1·232*
150	29·72†	1·278	(1·280)
−273	20·28†	0·873	0·87*

* Extrapolated from above melting point of polyethylene.

† Extrapolated from v_2 at 20° and 100°C.

( ) Value observed directly.

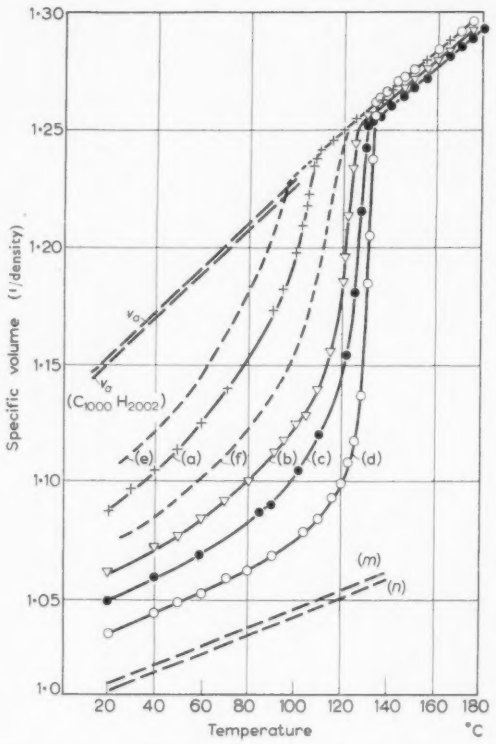


FIG. 2. Specific volume/temperature relationship for some polyethylenes.

(a) Alkathene 7, (b) Alkathene H.D., (c) Hostalen F, (d) Marlex 50, (e) (f) Values published by HUNTER and OAKES.

v_a = calculated volume of liquid paraffins for $(CH_2)_\infty$ or (C_{1000} H_{2002};

(*m*),(*n*) = calculated volume of crystalline material v_c with $\alpha = 4 \times 10^{-4}$.

m; $v_c = 1{\cdot}014$ at 20°C.

n; $v_c = 1{\cdot}010$ at 20°C.

RESULTS

The specific volume of various types of polyethylene as a function of temperature are shown in Fig. 2. The calculated volumes v_c and v_a are represented by broken lines; the values of v_a for $(CH_2)_\infty$ and for $C_{1000}H_{2002}$ calculated as shown in the Appendix are indicated. Also included are previously published data on several other polyethylenes of varying molecular weight and degrees of branching. In each case, it is seen that at high temperatures, the points fall on approximately the same linear curve, which represents the thermal expansion of an entirely molten material. As the temperature is reduced, there is a rapid decrease in the specific volume due to partial crystallization ($v_c < v_a$). For low-pressure polyethylenes with low degrees of branching, the temperature at which

crystals first start to form lies close to the temperature (136·5°C) deduced by MANDELKERN *et al.*[15] for the melting point of an infinitely long linear polymethylene chain. For such polymers there is a very sudden decrease in the specific volume due to to the very narrow range of temperature within which most of the crystals are formed. In high-pressure polyethylene, on the other hand, crystals begin to form only at temperatures below about 115°C and the decrease in specific volume is far less rapid, crystallization occurring down to temperatures of the order of 60°C or less. This behaviour is to be expected from the presence of side branches which limit the size of individual crystals and thereby reduce their size and melting point.

From these curves, the percentage of crystalline material at any temperature can be calculated using the corresponding values of v_a and v_c. Some typical curves are shown in Fig. 3. These illustrate the extensive range of melting of high-pressure

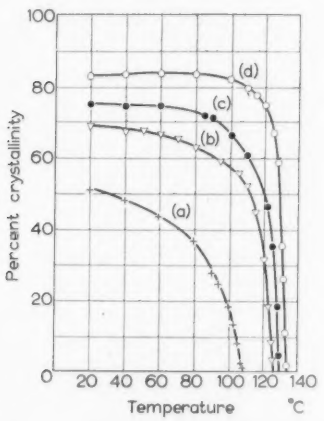


FIG. 3. Percentage crystallinity as a function of temperature. (a) Alkathene 7, (b) Alkathene H.D., (c) Hostalen F, (d) Marlex 50. $v_c = 1{\cdot}0146$ (a); $1{\cdot}0146$ (b); $1{\cdot}0114$ (c); $1{\cdot}0103$ (d). $\alpha = 4 \times 10^{-4}$.

branched polyethylene and the much smaller melting range of the low-pressure material.

The values derived for the percentage of crystallinity are based on the assumption that the volumes of the crystalline and amorphous regions are additive. They depend, moreover, on the values chosen for the specific volumes v_c and v_a, and on their variation with temperature. In his measurement of the crystallinity of various polyethylenes, measured both by X-rays and by nuclear resonance, SMITH,[2] found a linear relationship between crystallinity and specific volume, thereby implicitly confirming the additivity of amorphous and crystalline volumes. ASHWORTH[16] observed no change in unit-cell volume on straining a polyethylene specimen, so that the presence of amorphous chains attached to each crystal is unlikely to have an appreciable effect on its specific volume. In their studies on the relationship of branching and crystallinity, REDING and LOVELL[17] chose values of 1·00 and 0·850 for the density at 25°C of the crystalline and amorphous regions, respectively. Although these values differ from ours, this has little effect on the percentage of crystallinity deduced from our observed density for Alkathene: 49 per cent as against our calculated value of 50 per cent.

A small error in the coefficient of expansion of the unit cell may lead to erroneous figures for the degree of crystallinity, particularly at intermediate temperatures (about 50°C). Values of between 3·5 and 4×10^{-4} give approximately constant values for the crystallinity of Marlex up to some 60°C.

The highest degree of crystallinity, associated with the highest density, is observed in Marlex 50. At 20°C we deduce a value of 83 per cent crystallinity, as against 93 per cent obtained by SMITH[2] who used X-ray and nuclear-resonance techniques. The published density of Marlex (0·96) is difficult to reconcile with as high a degree of crystallinity as 93 per cent and would lead to an extrapolated value of 0·97 for the density of completely crystalline polyethylene—a value at variance with the measured unit-cell dimensions. Using the BRYANT[18] relation between density and crystallinity (which is most reliable in this region), a value of about 80 per cent is deduced, in fair agreement with our estimate of 83 per cent. If, from SMITH's values for specific volume, a figure of 0·04 is substracted, values for v_c and v_a (1·01 and 1·16) in good agreement with most published values are obtained.

EQUIVALENT CRYSTAL SIZE DISTRIBUTION

The intensity distribution of the X-ray diffraction pattern of polyethylene is insufficient to provide accurate estimates of the individual crystal sizes. An alternative approach which may be useful

attempts to express the observed melting data in terms of an equivalent paraffin length. Crystals of normal paraffins melt at a temperature which depends on their length, i.e., on the number of carbon atoms per molecule. Fig. 4 shows the relationship between the melting point of a paraffin (C_nH_{2n+2}) and the reciprocal ($1/n$) of the number of carbon atoms per molecule, i.e., the concentration of molecular ends. The maximum value of approximately 136·5°C corresponds closely to the melting point of low-pressure unbranched polyethylene of high molecular weight. If the temperature of a polyethylene specimen is increased by a small amount, the resultant decrease in crystallinity may be ascribed to the preferential melting of the smallest crystals present at that temperature. We may compare the size of these crystals with those of normal paraffin molecules C_nH_{2n+2} which melt in the same temperature range. It is not assumed that the crystals in polyethylene are necessarily equal in length to paraffin molecules melting in the same temperature range, since the conditions for melting are not identical; for example, no account is taken of possibly strained regions near the surface of crystals. Nevertheless, this comparison offers a convenient means of describing the size distribution of the crystallite regions of polyethylene in a simple form. In practice, the decrease in crystallinity over any small temperature range is compared with the difference in the value of n for normal paraffins melting at the two extremes of the temperature range covered. The crystallinity data of Fig. 3 have been converted to the n-paraffin distributions shown in Fig. 5, which indicates not only the difference in average equivalent crystal size as between various types of polyethylene, but also gives some idea of the distribution of these equivalent sizes, measured along the molecular chain length. No information is available on the lateral dimensions of the crystals, which have a far less-marked influence on its melting properties. The dimensions obtained in this way appear to be of the correct order of magnitude, i.e., 100–250 Å. They are somewhat smaller than the values deduced by BUNN[11] from X-ray data and by BUECHE[19] from purely theoretical considerations, but the differences are not sufficiently large to be significant at this stage. While it is not implied that the crystals in polyethylene have in fact this size distribution, the concept of an equivalent paraffin crystal is of value in comparing different distributions between various types of polyethylene. Highly branched polyethylene of low melting point would have a distribution consisting of small crystal sizes, whereas a hard material containing predominantly large crystals would be represented by

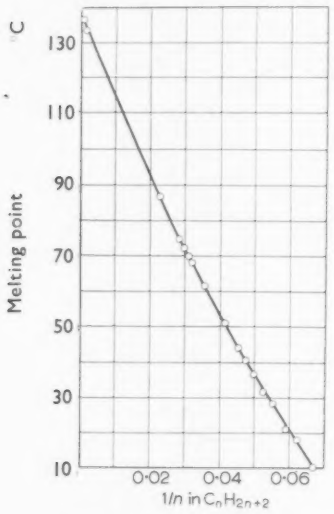


FIG. 4. Melting point of n-paraffins C_nH_{2n+2} versus $1/n$.

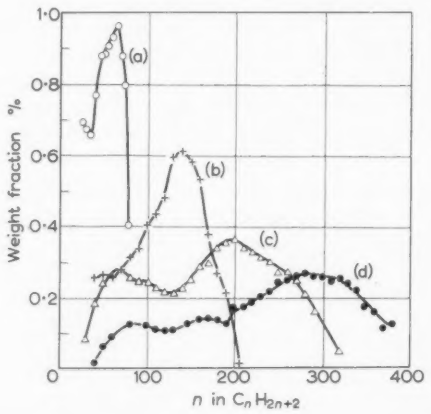


FIG. 5. Crystal size distribution of polyethylenes (in terms of equivalent n-paraffins). (a) Alkathene 7, (b) Alkathene H.D., (c) Hostalen F, (d) Marlex 50.

VOL. 4 1958

a distribution confined almost entirely to large values of n.

Mixtures of polyethylenes

The value of representing a polyethylene by its equivalent crystal size distribution is best illustrated by considering the results obtained from a mixture of two polyethylenes of widely differing characteristics. The two materials considered were a highly branched Alkathene 7 and a very lightly branched Marlex 50. Specimens of these, free of antioxidant, were degassed, and their density measured over the temperature range shown in Fig. 6. Temperature changes were made very slowly to allow thermal equilibrium to be reached, but the values obtained differed little from identical specimens measured previously in which temperature changes were far more rapid. Measurements were made on Marlex 50 provided both in the form of a thin film and as a solid block, the differences in specific volume being only about 0·001.

A fifty–fifty mixture by weight was prepared from the two materials by protracted milling. The specific volume temperature curve is shown in Fig. 6. From these data the crystallinity was deduced as a function of temperature and of equivalent paraffin length (Fig. 7). It is seen that Alkathene 7 shows a rapid drop in crystallinity from a temperature of 20°C onwards (or an equivalent paraffin C_n of 16), while for Marlex 50 crystallinity is constant to within about 1 per cent up to a temperature of 80–90°C, or an equivalent n-value of about 40 or 50. (The very slight curvature observed below this temperature depends on the precise value chosen for the coefficient of expansion of the crystal lattice.) The broken curve represents the calculated curve assuming the direct addition of the crystallinity of Marlex to that of Alkathene. The figure shows that some form of interference in crystallinity takes place as a result of mixing.

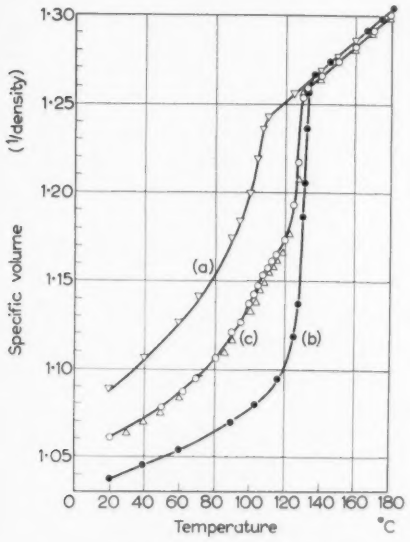


FIG. 6. Specific volume of polyethylene mixes. (a) Alkathene 7, (b) Marlex 50, (c) o Alkathene 7 (50 per cent) + Marlex 50 (50 per cent) after milling and melting △ Original mix after remilling.

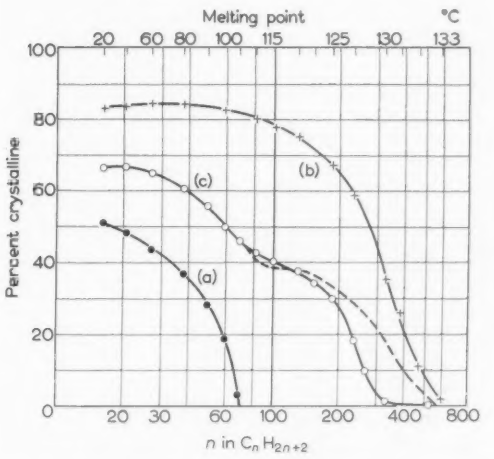


FIG. 7. Crystallinity of polyethylene mixes as a function of temperature and equivalent paraffin length. (a) Alkathene 7, (b) Marlex 50, (c) Alkathene 7 (50 per cent) + Marlex 50 (50 per cent) after milling and melting; (c) - - - - calculated on the assumption of a simple mixture of Alkathene 7/Marlex 50.

The results of an analysis in terms of equivalent n-paraffins is shown in Fig. 8. For high values of n (about 250), the figures depend on the precise value chosen for the melting point for an infinite paraffin (here taken as 136·5°C), but apart from a change in scale this does not affect the shape of the curves to any marked extent.

The curves for Alkathene 7 and for Marlex 50

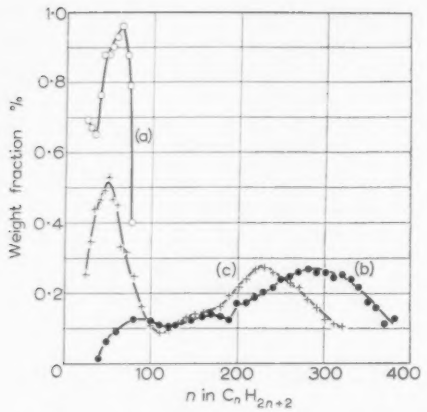


FIG. 8. Crystal size distribution of polyethylene mixes. (a) Alkathene 7, (b) Marlex 50, and (c) Alkathene 7 (50 per cent) + Marlex 50 (50 per cent) after milling and melting.

each show peaks, at $n = 60$ and 290, respectively; the mixed polyethylene shows two peaks, at 50 and 230 approximately. The milling process (and subsequent melting during the determination of density) has not produced a homogeneous material, and the mechanical properties would be expected to show two distinct transitions, at the temperatures corresponding to the melting of C_{50} and C_{230}, i.e. 90° and 127°C.

Effect of branching

The persistance of characteristic crystal sizes, even in a mixed polyethylene, is an indication of the permanent nature of the irregularities (branching, unsaturation, and finite molecular weight) which may limit the crystal size. Of these, branching is certainly the most frequent, at least in Alkathene 7. The crystallinity-branching data published by REDING and LOVELL[17] (which can be translated to density-branching) indicates about 1·3 ethyl branches per 100 carbon atoms in Alkathene 7. For a somewhat similar high-pressure polyethylene (DYNH), SMITH[2] found 2·15 branches with methyl ends per 100 C; of these 1·44 are ethyl branches. Assuming an approximate value of 1·5 branches per 100 C for these low-density polyethylenes, there would be an average of about 67 C between successive branches. For Alkathene 7 this is also the average length of the maximum equivalent *n*-paraffin (Fig. 5). If branching points are distributed at random, only a small fraction of the chains between branching points would be considerably longer than 67 C, so that if crystals are limited by such branch points, this would explain the rapid decrease in equivalent crystal size beyond this length. For Marlex 50 polyethylene, the degree of branching is very low, below one methyl per 300 carbon atoms, and the explanation of branching cannot be advanced to account for the limited crystal size. To some extent crystallinity will also be limited by unsaturation and molecular chain ends, but the major restriction may be expected to arise from the characteristics of crystal growth such as apply to the crystallization of low-molecular-weight compounds. With increased perfection of structure, and a reduced degree of amorphous material, the influence of a single structural defect will be greatly enhanced.

CONCLUSIONS

Measurements on the crystallinity of various types of polyethylene have been compared, and show that the main difference between high-density polyethylene (such as Alkathene) and high-density polyethylene (Ziegler or Marlex) arises from the average size and size distribution of the crystalline regions. In Alkathene 7, which is highly branched, the length of these crystals is limited by the number of side branches which interfere with the regularity of structure. The distribution of crystallite sizes is, moreover, much wider for Alkathene, so that there is a wide range of melting temperatures. In Marlex and Ziegler polyethylene, the individual crystallites are much bigger, and their melting point approaches that of an infinitely long polymethylene material. It would therefore appear that little improvement is to be expected in the temperature properties of polyethylene.

The distribution of crystal sizes can be expressed in terms of an equivalent paraffin distribution, i.e., a range of *n*-paraffins whose melting points would parallel that observed by a polyethylene.

The concept of equivalent crystal size offers a convenient method characterizing the melting properties of a series of polyethylene specimens of widely differing degrees of branching. Although there appears to be some relationship between equivalent crystal size and length of unbranched

VOL. 4 1958

chain, it cannot be assumed that the former does in fact represent the distribution in length of the crystals present.

When polyethylenes of different degrees of branching are melted together, the distribution of equivalent crystal size is related to each of the two components, and there is only a limited amount of interaction between them. The determination of equivalent crystal size distribution may therefore serve as a useful method of characterizing various polyethylenes.

Acknowledgements—The authors wish to thank Mr. D. R. ASHWORTH for X-ray data on a number of polyethylenes, and the Chairman of Tube Investments, Ltd., for permission to publish this paper.

Appendix: Specific volume of amorphous *n*-paraffins

The specific volume (reciprocal of density ρ) of a linear paraffin C_nH_{2n+2} will depend on n mainly owing to the relative proportions of methylene ($-CH_2-$) and methyl ($-CH_3$) groups. If at any temperature the volumes (in Å^3) occupied by $-CH_2-$ and $-CH_3$ groups can be taken as v_2 and v_3, respectively, and these volumes are additive, the specific volume of C_nH_{2n+2}

$$\frac{1}{\rho} = \frac{[(n-2)v_2+2v_3]10^{-24}}{[(n-2)14{\cdot}016+2\times 15{\cdot}024]1{\cdot}66035\times 10^{-24}}$$

$$= \frac{v_2+(2/n)(v_3-v_2)}{14{\cdot}016+(2{\cdot}016/n)}\times 0{\cdot}60228.$$

A plot of $[14{\cdot}016+(2{\cdot}016/n)]/\rho$ against $1/n$ should therefore give a straight line of slope $2\times 6{\cdot}0228$ (v_3-v_2) and of intercept (at infinite molecular weight, when $1/n = 0$) of $0{\cdot}60228v_2$.

The values in Fig. 9, calculated from published values of the density of normal paraffins at 20° and 100°C, give the required linear plot over the range C_5H_{12} to $C_{16}H_{34}$ at 20°C (higher paraffins are crystalline) and over the range $C_{11}H_{34}$ to $C_{64}H_{130}$ at 100°C. The values of v_2 and v_3 deduced are 26·81 and 54·37 at 20°C and 28·6 and 61·4 at 100°C. The corresponding specific volumes for infinite *n*-paraffins would then be 1·152 and 1·229 at these temperatures. A linear extrapolation to 0°K leads to a specific volume for an infinite linear paraffin of 0·873. The volume expansion, assuming that volume expansion is uniform with temperature, is

$$2{\cdot}26\times 10^{-26}\times 6{\cdot}0228\times 10^{23}/14{\cdot}016$$

$$= 0{\cdot}971\times 10^{-3}\text{cm}^3/\text{g per °C}.$$

The calculated volume is 1·248 at 120°C and 1·277 at 150°C (as compared with observed values for molten unirradiated polyethylene at the latter temperature of 1·282±0·003).

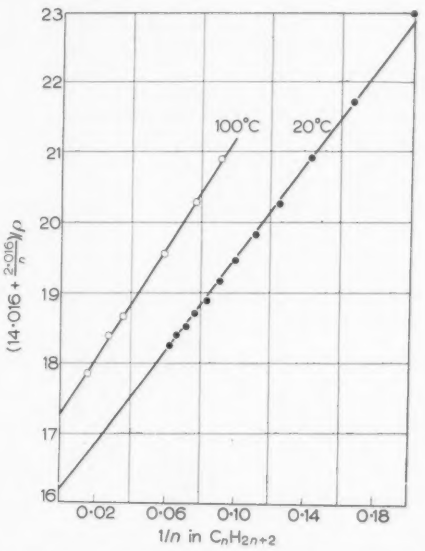


FIG. 9. Volume of liquid *n*-paraffins at 20 and 100°C.

Minor discrepancies are to be expected from the presence of side chains and of the finite molecular weight in polyethylenes. For example, the effect of reducing the average molecular weight to 14,000 ($n = 0{\cdot}001$) would in itself increase the calculated specific volume by 0·0023 at 20°C. To this must be added the effect of the side chains present in all polyethylenes, a correction which has not been calculated. The differences between the specific volume of amorphous polyethylene and the calculated value for an infinite linear paraffin lie well within the expected accuracy of the simple calculation.

REFERENCES

1. HARLEN F., SIMPSON W., WADDINGTON F. B., WALDRON J. D. and BASKETT A. C. *J. Polymer Sci.* **18,** 589 (1955).
2. SMITH D. C. *Industr. Engrg. Chem.* (Industr.) **48,** 1161 (1956).
3. KRIMM S. and TOBOLSKY A. V. *J. Polymer Sci.* **7,** 57 (1951).
   HERMANS P. H. and WEIDINGER A. *J. Polymer Sci.* **4,** 709 (1949).
   BUNN C. W. and ALCOCK T. C. *Trans. Faraday Soc.* **41,** 317 (1945).
4. RICHARDS R. B. *J. Appl. Chem.* **1,** 370 (1951).

5. RAINE H. C., RICHARDS R. B. and RYDER H. *Trans. Faraday Soc.* **41,** 56 (1945).
DOLE M., HETTINGER W. P., LARSON N. R. and WETHINGTON J. A. *J. Chem. Phys.* **20,** 781 (1952).
6. HUNTER E. and OAKES W. G. *Trans. Faraday Soc.* **41,** 49 (1945).
7. CHARLESBY A. *Proc. Phys. Soc.* **57,** 496 (1945).
8. CHARLESBY A. and ROSS M. *Proc. Roy. Soc.* A **217,** 122 (1953).
9. *Handbook of Chemistry and Physics* (36th Ed.) p. 1941. Chemical Rubber Publishing Co., Cleveland, Ohio (1954).
10. *Handbook of Chemistry and Physics* (36th Ed.) pp. 1964, 2065.
11. BUNN C. W. *Trans. Faraday Soc.* **35,** 482 (1939).
12. CHARLESBY A. *Proc. Phys. Soc.* **57,** 514 (1945).
13. WALTER E. R. and READING F. P. *J. Polymer Sci.* **21,** 561 (1956).
14. MATTHEWS J. L., PEISER H. S. and RICHARDS R. B. *Acta Cryst.* **2,** 85 (1949).
15. MANDELKERN L., HELLMANN M., BROWN D. W., ROBERTS D. E. and QUINN F. A., JR. *J. Amer. Chem. Soc.* **75,** 4093 (1953).
16. ASHWORTH D. R. Private communication.
17. READING R. P. and LOVELL C. M. *J. Polymer Sci.* **21,** 157 (1956).
18. BRYANT W. M. D., TORDELLA C. T. and PIERCE R. H. H. Communication to 118 Meeting Amer. Chem. Soc., Chicago (1950).
19. BUECHE F. *J. Polymer Sci.* **22** ,113 (1956).

VOL.
4
1958

BOOK REVIEW

Semiconductor Surface Physics. Edited by R. H. KINGSTON. University of Pennsylvania Press, 1957. xv+413 pp.

THE realization that the performance of semiconductor devices depends markedly on the properties of the surfaces has led to the development of a new branch of physics. Much progress has been made in the last four years, and the latest theoretical and experimental results were discussed at a conference on the Physics of Semiconductor Surfaces held in Philadelphia last year. The invited papers and some of the discussion have been assembled in book form and are now published under the title *Semiconductor Surface Physics*. The contributions, which are mostly reviews of particular aspects of the subject by active workers in the field, are divided into four sections: Clean Surfaces, Real Surfaces, Adsorption and Catalysis and Oxidation. In the first section the precautions necessary to obtain clean surfaces are detailed, and the electrical properties of such surfaces are discussed. The second section, occupying half the book, contains reports on mobility, conductance, recombination, storage, and photoeffects in germanium surfaces and a mention of lead sulphide. The remaining sections are of direct interest to the physical chemist; they include articles on gas reactions at surfaces, catalysis and semiconductor properties, and the oxidation of germanium and other metals.

In a collection of this kind the reader gains by having presented to him the ideas of many acknowledged experts, but he must be prepared to meet a few differences of opinion and some repetition of matter. It is unfortunate that the review of surface-state theories given by HERRING at the conference could not be included, since this, together with a summary of the main conclusions reached, would have added considerably to the value of the book.

The general standard of the contributions is high, the article on noise by MCWHORTER being particularly good, but one or two articles might be criticized as giving too much detail of the author's own experimental work and too little about the general techniques available.

The book is well produced, with clear diagrams and print and an adequate index, and it can be recommended.

C. HILSUM

BOOKS RECEIVED

HENISCH, H. K. Rectifying Semi-conductor Contacts (International Series of Monographs on Physics) 372 pp. (Clarendon Press: Oxford University Press, February 1957) 70s.

HARTREE, DOUGLAS R. The Calculation of Atomic Structures, 1957, 181 pp. (John Wiley & Sons Inc., New York: Chapman & Hall Ltd., London 1957) 40s.

GRIFFITH, R. H. and MARCH, J. D. F. Contact Catalysis 299 pp. (Clarendon Press: Oxford University Press, June, 1957) 50s.

MOULON, J. M. Les transistrons dans les Amplificateurs 316 pp. (Gauthier-Villars, Paris, 1956).

SEITZ, F. and TURNBULL, D. (Editors) Solid State Physics, Vol. 4. Advances in Research and Application 540 pp. (Academic Press, New York, 1957) $12.00.

WOOSTER, W. A. Experimental Crystal Physics 115 pp. (Oxford University Press, 1957) 18s.

MCLEAN, D. Grain Boundaries in Metals 346 pp. (Oxford University Press, 1957) 50s.

TRIBALAT, S. Rhénium et Technétium, Monographies de Chimie Physique 172 pp. (Gauthier-Villars, New York, 1957).

GOLDMAN, E. J. (Editor) The Science of Engineering Materials 528 pp. (John Wiley & Sons Inc., 1957) 96s.

FIELD, F. H. and FRANKLIN, J. C. Electron Impact Phenomena and the Properties of Gaseous Ions, Vol. I, 349 pp. (Academic Press, New York, 1957) $8.50.

NEUMAN, W. F. and NEUMAN, M. W. The Chemical Dynamics of Bone Mineral 209 pp. (University of Chicago Press, 1958) $5.00.

EFRON, A. Basic Physics, Vols 1–2, 724 pp. (John F. Rider, Inc., 1957) $7.60.

SEITZ, F. and TURNBULL, D. (Editors) Solid State Physics, Vol. 5, 455 pp. (Academic Press, New York, 1957) $11.00.

ERRATUM

ABRIKOSOV A. A. The magnetic properties of superconducting alloys. *J. Phys. Chem. Solids* **2,** 199 (1957).

The Editor regrets that in the process of translating the above article numerous errors were introduced into the text as well as the formulae. Since these errors were too numerous to list in detail, the reader is referred to the article "The magnetic properties of superconductors of the second group" by the same author, published in *Zh. eksp. teor. fiz., SSSR.* **32,** 1442 (1957).

VOL.
4
1958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 241–255.

A THERMODYNAMIC THEORY OF "WEAK" FERROMAGNETISM OF ANTIFERROMAGNETICS

I. DZYALOSHINSKY

Institute for Physical Problems, Academy of Sciences of the U.S.S.R., Moscow

(*Received* 19 *February* 1957)

Abstract—A thermodynamic theory of "weak" ferromagnetism of α-Fe_2O_3, $MnCO_3$ and $CoCO_3$ is developed on the basis of LANDAU's theory of phase transitions of the second kind. It is shown that the "weak" ferromagnetism is due to the relativistic spin-lattice and the magnetic dipole interactions. A strong dependence of the properties of "weak" ferromagnetics on the magnetic crystalline symmetry is noted and the behaviour of these ferromagnetics in a magnetic field is studied.

CERTAIN antiferromagnetic crystals, e.g. α-Fe_2O_3 and the carbonates $MnCO_3$ and $CoCO_3$, exhibit spontaneous magnetization. The magnitude of their magnetic moment is, however, extremely small and varies between the limits of 10^{-2}—10^{-5} of the nominal value. In α-Fe_2O_3 it represents 0·02 per cent according to NÉEL and PAUTHENET's measurements[1] and in the carbonates, according to BOROVIK-ROMANOV and ORLOVA,[2] 0·2 per cent for $MnCO_3$, and $\simeq$ 2 per cent for $CoCO_3$. The smallness of the spontaneous moments indicates that we are not dealing here with the usual exchange ferromagnetism when, except in the vicinity of the Curie point, the spontaneous moment is of the same order of magnitude as the nominal value.

In a recent work of YIN-YUAN LI[3] this effect is explained by antiferromagnetic domains with magnetized walls. The author himself points out that the formation of such domains is not energetically advantageous and that it would be possible only due to non-uniformities and impurities in a crystal. These ferromagnetic properties of the crystal would then depend strongly on its purity and uniformity, they would vary from sample to sample and, would disappear completely in an ideal crystal.

The work reported in reference 2 was carried out with very pure crystals and showed that ferromagnetism is still retained and therefore the connexion between the weak ferromagnetism and any non-uniformities seems very unlikely. It will be shown that many properties of the "weak" ferromagnetics follow directly from thermodynamics and may be fully explained in terms LANDAU's theory[4,5] of phase transitions of the second kind by their magnetic symmetry.

MAGNETIC SYMMETRY

Besides the smallness of the spontaneous magnetic moment the "weak" ferromagnetism is very sensitive to the crystal symmetry.

We shall consider two isomorphous antiferromagnetics α-Fe_2O_3 and Cr_2O_3. Both belong to the

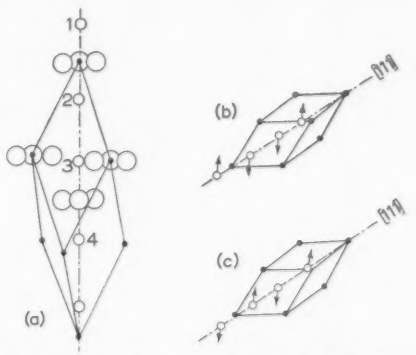


FIG. 1. (a)—Unit cells of α-Fe_2O_3 and Cr_2O_3: The large circles represent the oxygen ions, the small ones ferric ions. (b)—Antiferromagnetic structure of α-Fe_2O_3. (c)—Antiferromagnetic structure of Cr_2O_3.

rhombohedral system and the distribution of their atoms possesses D_{3d}^6 space group symmetry.[6] In their unit cell there are four Fe^{3+} or Cr^{3+} ions, distributed along the space diagonal of the rhombohedron (Fig. 1). Neutron diffraction studies of SHULL *et al.*[7] and BROCKHOUSE[8] showed that in the antiferromagnetic state the magnetic unit cell is identical with the space one. Spins of the ions 1, 2, 3 and 4 differ in sign only and their sum in the cell is equal to zero, and in α-Fe_2O_3 $s_1 = -s_2 = -s_3 = s_4$ while in Cr_2O_3 $s_1 = -s_2 = s_3 = -s_4$ (Fig. 1a and b). Direction of the spins with respect to the crystalline axes was not determined for Cr_2O_3. Depending on temperature α-Fe_2O_3 may occur in two different antiferromagnetic states: at $T < 250°K$ the spins are along the crystal axis (state I), and at $250°K < T < 950°K$ they lie in one of the vertical planes of symmetry making a small angle with the basal plane (111) (state II).

It appears that Cr_2O_3 is not ferromagnetic at all and α-Fe_2O_3 is ferromagnetic only in the state II. When, with lowering of temperature, α-Fe_2O_3 changes from the II state into the I state its spontaneous magnetic moment disappears.

Let us assume that we know the symmetry of the antiferromagnetic distribution of the spins when their sum in the unit cell is equal to zero. A question then arises: is ferromagnetism possible with this symmetry? In other words, is it possible to change the spin direction or magnitude, without departing from the symmetry of the original distribution, in such a way as to make the sum of the spins in the unit cell non-zero in value?

Thus, for example, in the II state α-Fe_2O_3 ferromagnetism would arise if all spins leave the symmetry plane and rotate towards one another (Fig. 2). The possibility of such a rotation depends

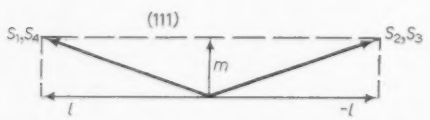


FIG. 2. Projection of ion spins on (1) for α-Fe_2O_3.

on whether the magnetic symmetry of the II state and the symmetry with the spins rotated are identical.

The symmetry of a magnetic crystal is determined not only by the distribution of its atoms but also by the average value of the spin at every point $s(x, y, z)$. It is known[5] that the vector quantity $s(x, y, z)$ may possess in addition to the usual symmetry transformations, i.e. axes of rotation, symmetry planes and translations, also a symmetry element R consisting of the change of the spin sign $s(x, y, z) \to -s(x, y, z)$. This is due to the invariance of the equations of mechanics with respect to simultaneous change of the sign of time and the sign of magnetic fields and spins.

If the spin distribution possesses R by itself, then $s(x, y, z) = -s(x, y, z) = 0$ and the crystal is paramagnetic. A non-zero value of $s(x, y, z)$ may be, however, invariant with respect to various combinations of R with axes of rotation, planes of symmetry and translations. Thus in addition to the 230 known space groups, which describe all the possible types of symmetry of the distribution of atoms in crystals, new "magnetic" space groups arise which describe symmetry of the spin distribution.

All groups obtainable in this way have not yet been constructed. However, to describe the symmetry of a ferromagnetic body we need not know the corresponding space group. The symmetry of the ferromagnetic and, in general, all macroscopic, properties is determined by the symmetry class, i.e. by the point group obtained from the corresponding space group by replacement of translations by identical transformations, and helical axes and glide surfaces by rectilinear axes and planes. All the possible "magnetic" classes were described by TAVGER and ZAITSEV.[9] It was found that, in addition to the usual 32 classes which possess the element R by itself, there are also 58 special "magnetic" classes.

We shall find the classes of symmetry of α-Fe_2O_3 and Cr_2O_3 in the antiferromagnetic state. As shown earlier, α-Fe_2O_3 may occur in two antiferromagnetic states: with spins directed along the crystal axis (state I) and with spins lying in one of the planes of symmetry (state II). From considerations of symmetry we should also discuss another possible spin orientation, along one of the second-order axes (state III).

Using the description of the space group D_{3d}^6 given in reference 10 we can easily find to which classes the states I, II and III of α-Fe_2O_3

VOL. 4 1958

belong. We shall now write down the elements of symmetry which form the respective classes:*

$$C_3,\ U_2,\ I,\ S_6,\ \sigma_d \qquad \text{(I)}$$

$$U_2,\ I,\ \sigma_d \qquad \text{(II)}$$

$$U_2R,\ I,\ \sigma_dR. \qquad \text{(III)}$$

A spontaneous magnetic moment $\boldsymbol{m}$ may exist only when it is invariant under the action of all transformations of the symmetry of the class. In the I state $\boldsymbol{m}$ is zero since the corresponding class contains a third-order axis and a second-order one perpendicular to it. In the states II and III a non-zero spontaneous moment may exist and it is directed along a second-order axis in the II state and in the III state it lies in a plane of symmetry.

The experimental observation that α-Fe_2O_3 is ferromagnetic in the II state and not in the I state is a direct consequence of the magnetic symmetry. Symmetry considerations make it also possible to determine the direction of the spontaneous magnetic moment. In this case it is directed along that one of the second-order axes, perpendicular to a plane of symmetry, in which the antiferromagnetic components of the spins lie.

We shall now consider Cr_2O_3. As shown earlier, spins of the Cr^{3+} ions differ only in their sign, and $\boldsymbol{s}_1 = -\boldsymbol{s}_2 = \boldsymbol{s}_3 = -\boldsymbol{s}_4$. Such an antiferromagnetic structure possesses element of symmetry IR. It therefore follows that Cr_2O_3 cannot be ferromagnetic at all. We find on applying the element IR to the magnetic moment vector $\boldsymbol{m}$ that

$$\boldsymbol{m} = -\boldsymbol{m} = 0.$$

The above considerations make it possible to decide whether ferromagnetism is possible in a given antiferromagnetic structure of a crystal and to find directions of a spontaneous magnetic moment that are permissible in given structural symmetry. The symmetry does not limit in any way the magnitude of the magnetic moment. For example, in α-Fe_2O_3 in the II state it is possible, without affecting the symmetry, to rotate the spins about the crystal axis to make their directions coincide and the crystal would then be a pure ferromagnetic. To explain the smallness of the spontaneous magnetization and to show the close relationship between the "weak" ferromagnetism and antiferromagnetism it is necessary to apply the thermodynamic theory of transitions of the second kind.

* Here C_3 denotes a vertical axis of third order, U_2 = a horizontal axis of second order, σ_d = a vertical plane of symmetry, S_6 = a mirror-rotation axis of sixth order and I = an inversion operation.

LANDAU'S THEORY OF PHASE TRANSITIONS OF THE SECOND KIND[4,5]

The state of a magnetic crystal is described by the average density of the spin distribution $\boldsymbol{s}(x, y, z)$. If in a unit cell there are several magnetic ions which cannot be reduced by translations then the average spin of each type of ions is described by its own function $\boldsymbol{s}_a(x, y, z)$. The symmetry of a magnetic crystal is identical with the symmetry of the vector function $\boldsymbol{s}(x, y, z)$ and is given by one of the "magnetic" space groups.† In a phase transition of the second kind the spin density $\boldsymbol{s}(x, y, z)$, i.e. the magnitude and the orientation of the spins in the crystal, varies continuously but its symmetry changes suddenly, and at the transition point itself the symmetries of both phases are identical.

We shall discuss the transition from a paramagnetic into an antiferromagnetic state. In the paramagnetic phase $\boldsymbol{s}(x, y, z) = 0$ and its "magnetic" symmetry is described by one of the 230 "magnetic" groups with the element R by itself.‡ In the antiferromagnetic phase $\boldsymbol{s}(x, y, z) \neq 0$ (but the integral $\int \boldsymbol{s} dV$ of the spin density in the unit cell is equal to zero) and its symmetry is the same as the symmetry of the $\boldsymbol{s}(x, y, z)$ function.

From group theory it is known that any function can be represented as a linear combination of functions which transform one into another under all transformations of the group. Let $\phi_1(x, y, z)$, $\phi_2(x, y, z)$, . . . transform one into another under all transformations of the symmetry group of the paramagnetic phase. Then the spin density in the antiferromagnetic phase may be represented by

$$s_a(x,y,z) = \sum_i a_{ai} c_i \phi_i(x,y,z)$$

† The atomic distribution symmetry can be obtained from the average spin distribution symmetry if in the appropriate magnetic group the element R is regarded as an identical transformation.

‡ If the usual space group of the crystal is G, then in the paramagnetic state its magnetic group is $G \times R$.

where $a_{\alpha i}$ are temperature- and pressure-independent coefficients chosen so that the components of the spin density $s_\alpha(x, y, z)$ transform as pseudovector components.

In this expansion instead of the functions ϕ_i we may of course use any linear combination of them. It is always possible to select ϕ_i so that they can be resolved into a series of sets, containing a minimum number of functions, such that in all transformations of the symmetry group the functions of each set transform one into another. Each such set is an irreducible representation of the symmetry group. In terms of these functions the expansion of $s(x, y, z)$ has the form

$$s_\alpha(x,y,z) = \sum_n \sum_i a_{\alpha i}{}^{(n)} c_i{}^{(n)} \phi_i{}^{(n)}(x,y,z) \qquad (1)$$

where n = the number of an irreducible representation and i = the number of a function in its base.

The symmetry group of a paramagnetic crystal contains R by itself. Since $R^2 \equiv 1$ and R is interchangeable with all axes of rotation, planes of symmetry and translations, therefore all the irreducible representations of the group can be divided into two types of representations: the first type with functions which do not change under the action of R, and those of the second type with functions which change their sign. Since $s(x, y, z)$ itself changes its sign under the action of R, the above expansion would contain only those functions which change their sign under the action of R.

The thermodynamic potential Φ of a crystal with spin density $s(x, y, z)$ depends on temperature, pressure and the coefficients $c_i{}^{(n)}$. Dependence of $c_i{}^{(n)}$ themselves on p,T, and therefore the symmetry of the crystal in the antiferromagnetic phase are determined thermodynamically from the condition of minimum of Φ.

In the paramagnetic phase all $c_i{}^{(n)}$ are equal to zero. Since the change of state of a crystal at a phase transition of the second kind is continuous therefore the coefficients $c_i{}^{(n)}$ become zero at the phase-transition point, and they have very small values in the antiferromagnetic phase near the transition point. It follows that near the transition point the potential Φ can be expanded into a power series of $c_i{}^{(n)}$.

In the expansion (1) we regarded $c_i{}^{(n)}$ as constant and $\phi_i{}^{(n)}(x, y, z)$ as changeable under transformations of the group. We can equally well regard $\phi_i{}^{(n)}$ to be constant and $c_i{}^{(n)}$ transformable in the same way as $\phi_i{}^{(n)}$ earlier. Since Φ is invariant under transformations of the symmetry, its expansion in terms of $c_i{}^{(n)}$ should contain only invariant combinations of the latter.

As shown earlier, all $c_i{}^{(n)}$ change their sign under the R transformation, and therefore in the expansion of Φ we can only have invariants composed of even powers of $c_i{}^{(n)}$. From group theory we know that from quantities transforming in irreducible representation we can form only one invariant of the second order which is a positive definite quadratic form always expressible as a sum of squares.

Consequently, retaining only the square terms, the expansion of the thermodynamic potential Φ has the form

$$\Phi = \Phi_0 + \sum_n A^{(n)} \sum_i c_i{}^{(n)2}$$

where $A^{(n)}$ are functions of p and T, and Φ_0 is the potential of the paramagnetic phase.

In the paramagnetic phase when Φ is minimum $c_i{}^{(n)} = 0$ and consequently all $A^{(n)} > 0$. Transition to the antiferromagnetic phase occurs when one of the functions $A^{(n)}$ changes its sign (at the transition point itself $A^{(n)} = 0$). Then $c_i{}^{(n)}$ of the corresponding irreducible representation are no longer zero. The crystal symmetry now changes and is given by the symmetry of the average spin density

$$s_\alpha(x, y, z) = \sum_i a_{\alpha i}{}^{(n)} c_i{}^{(n)} \phi_i{}^{(n)}(x, y, z).$$

To find the actual form of the symmetry we must take into account terms with the fourth power of $c_i{}^{(n)}$. Introducing

$$\eta^2 = \sum_i c_i{}^2, \; c_i = \gamma_i \eta$$

(we are omitting index n for the sake of brevity) we can write the expansion of Φ as

$$\Phi = \Phi_0(p, T) + \eta^2 A(p, T) + \eta^4 \sum_k B_k(p, T) f_k{}^{(4)}(\gamma_i)$$

where $f_k{}^{(4)}(\gamma_i)$ are invariants of the fourth order which can be made up of γ_i.

On finding γ_i from the minimum condition for Φ we can write Φ in the form usual in the theory of phase transitions of the second kind:

$$\Phi = \Phi_0(p,T) + \eta^2 A(p,T) + \eta^4 B(p,T)$$

VOL. 4 1958

from which in the usual way we find the dependence of η^2 on p and T.

The quantities η^2 and γ_i found in this way determine the symmetry of the average spin density $s_\alpha = \eta^2 \sum_i a_{\alpha i}\gamma_i\phi_i$ and, simultaneously, the crystal symmetry in the antiferromagnetic phase.

The magnetic transition in α-Fe_2O_3 occurs without a change of the unit cell. In this case the theory put forward above is greatly simplified. Now the average spin density $s_a(x, y, z)$ of each of the four ions of Fe^{3+} or Cr^{3+} (Fig. 1), since we are interested only in changes of the symmetry, may be regarded to be independent of co-ordinates, and the magnetic symmetry of the crystal is determined simply by the average values of the spins of each of the ions: $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, $\boldsymbol{s}_3$, $\boldsymbol{s}_4$. In the expansion (1) there is no need now to distinguish the coefficients $c_i^{(n)}$ and the functions $\phi_i^{(n)}$, and by their product $c_i^{(n)}\phi_i^{(n)}$ we now understand a linear combination of the components of $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, $\boldsymbol{s}_3$, $\boldsymbol{s}_4$ transforming according to the corresponding irreducible representation of the crystal symmetry group in the paramagnetic phase. For α-Fe_2O_3 and Cr_2O_3 this group consists of all the elements of symmetry of their space group D_{3d}^6, the transformation R and all possible combinations of R with transformations of D_{3d}^6. Such a group is called a direct product of D_{3d}^6 and R and is usually written $D_{3d}^6 \times R$.

Near the transition point the average ionic spins $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, $\boldsymbol{s}_3$, $\boldsymbol{s}_4$ are small and the thermodynamic potential Φ may be expanded as a power series of their components. As before only the even-powered terms enter the expansion since Φ should be invariant under the transformation R of replacement of all $\boldsymbol{s}$ by $-\boldsymbol{s}$. The expansion should be also invariant under all transformations of the symmetry of the space group D_{3d}^6 of the crystal. Since the transition occurs without a change of the unit cell all translations by a whole period of the lattice should be regarded as identical transformations. The group D_{3d}^6 goes over into a group which we shall call $\tilde{D}_{3d}^6$, isomorphous with the point group D_{3d} which determines the crystal class, and which differs from the D_{3d}^6 by replacement of certain rotation axes and symmetry planes by helical axes and glide surfaces. The distribution of the elements of symmetry of $\tilde{D}_{3d}^6$ in the crystal cell is given in.[10]

Axial vectors $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, $\boldsymbol{s}_3$, $\boldsymbol{s}_4$ form a reducible representation of the $\tilde{D}_{3d}^6$ group. To break it up into irreducible representations we shall introduce instead of $\boldsymbol{s}$ vectors $\boldsymbol{m}$, $\boldsymbol{l}_1$, $\boldsymbol{l}_2$, $\boldsymbol{l}_3$ given by

$$\begin{aligned} \boldsymbol{m} &= \boldsymbol{s}_1+\boldsymbol{s}_2+\boldsymbol{s}_3+\boldsymbol{s}_4 & \boldsymbol{s}_1 &= \tfrac{1}{4}(\boldsymbol{m}+\boldsymbol{l}_1+\boldsymbol{l}_2+\boldsymbol{l}_3) \\ \boldsymbol{l}_1 &= \boldsymbol{s}_1-\boldsymbol{s}_2-\boldsymbol{s}_3+\boldsymbol{s}_4 & \boldsymbol{s}_2 &= \tfrac{1}{4}(\boldsymbol{m}-\boldsymbol{l}_1-\boldsymbol{l}_2+\boldsymbol{l}_3) \\ \boldsymbol{l}_2 &= \boldsymbol{s}_1-\boldsymbol{s}_2+\boldsymbol{s}_3-\boldsymbol{s}_4 & \boldsymbol{s}_3 &= \tfrac{1}{4}(\boldsymbol{m}-\boldsymbol{l}_1+\boldsymbol{l}_2-\boldsymbol{l}_3) \\ \boldsymbol{l}_3 &= \boldsymbol{s}_1+\boldsymbol{s}_2-\boldsymbol{s}_3-\boldsymbol{s}_4 & \boldsymbol{s}_4 &= \tfrac{1}{4}(\boldsymbol{m}+\boldsymbol{l}_1-\boldsymbol{l}_2-\boldsymbol{l}_3) \end{aligned}$$

The vector $\boldsymbol{m}$ is simply the average magnetic moment of a unit cell.

In all transformations of the symmetry vectors $\boldsymbol{m}$ and $\boldsymbol{l}$ transform independently. The components m_z, l_{1z}, l_{2z}, l_{3z} transform according to unidimensional representations A_{2g}, A_{1g}, A_{1u}, A_{2u} of the point group D_{3d} (for tables of characters of irreducible representations see reference 11). The x- and y-components of the vectors $\boldsymbol{m}$ and $\boldsymbol{l}_1$ transform according to the same two-dimensions representation E_g and the same components of the vectors $\boldsymbol{l}_2$ and $\boldsymbol{l}_3$ according to E_u.

Knowing the irreducible representations according to which the components of $\boldsymbol{m}$ and $\boldsymbol{l}$ transform we can write down the most general form of the expansion permitted by the symmetry of the crystal. Limiting ourselves to the second-order terms we have for Φ:*

$$\left.\begin{aligned} \Phi = \frac{A_1}{2}l_1^2+\frac{A_2}{2}l_2^2+\frac{A_3}{2}l_3^2+\frac{B}{2}m^2+\frac{\alpha_1}{2}l_{1z}^2 \\ +\frac{\alpha_2}{2}l_{2z}^2+\frac{\alpha_3}{2}l_{3z}^2+\frac{b}{2}m_z^2+ \\ +\beta_1(l_{1x}m_y-l_{1y}m_x)+ \\ +\beta_2(l_{2x}l_{3y}-l_{2y}l_{3x}). \end{aligned}\right\} \quad (2)$$

In the expression (2) there are various magnetic energy terms corresponding to different types of the spin interaction. Terms of the type Al^2 do not depend on the orientation of the vector $\boldsymbol{l}$ and the spins $\boldsymbol{s}$ relative to the crystal axes, and they represent exchange interactions; terms of the type αl_z^2, $(l_x m_y - l_y m_x)$ arise from the relativistic spin-lattice interactions and from the magnetic dipole interaction and determine the magnetic

* We shall use a rectilinear system of co-ordinates with the z-axis directed along the crystal axis and the x-axis along one of the second-order axes.

anisotropy of the crystal. The ratio α/A of the magnetic isotropy energy to that of the exchange energy is proportional to the square of the ratio $(v/c)^2$ of the velocity of electrons in the crystal to the velocity of light and lies usually between 10^{-2} and 10^{-5}.

An important peculiarity of the crystals with the symmetry we are discussing now is the occurrence of "mixed" terms $l_{1x}m_y - l_{1y}m_x$ and $l_{2x}l_{3y} - l_{2y}l_{3x}$ in the expansion of the thermodynamic potential. From the mathematical point of view they arise because the x- and y-components of vectors $\boldsymbol{m}$ and $\boldsymbol{l}_1$ (or $\boldsymbol{l}_2$ and $\boldsymbol{l}_3$) are transformed according to the same representations. We shall see later that the first of the above "mixed" terms is responsible for the ferromagnetism of α-Fe_2O_3.

In the case of the paramagnetic phase all $A > 0$ and $B > 0$ and the state in which all $\boldsymbol{l}$ and $\boldsymbol{m}$ are zero corresponds to the minimum of the thermodynamic potential. The transition occurs at the point where one of A's becomes zero.* If the spin interaction in the crystal is purely of the exchange type, then below the transition point only the vector $\boldsymbol{l}$ which stands by the negative A is nonzero, and consequently in crystals with the symmetry considered here one of the following three types of antiferromagnetic structure occurs:

(1) $\boldsymbol{l}_1 \neq 0,\ \boldsymbol{l}_2 = \boldsymbol{l}_3 = \boldsymbol{m} = 0,\ \boldsymbol{s}_1 = -\boldsymbol{s}_2 = -\boldsymbol{s}_3 = \boldsymbol{s}_4$

which is the structure of α-Fe_2O_3;

(2) $\boldsymbol{l}_2 \neq 0,\ \boldsymbol{l}_1 = \boldsymbol{l}_3 = \boldsymbol{m} = 0,\ \boldsymbol{s}_1 = -\boldsymbol{s}_2 = \boldsymbol{s}_3 = -\boldsymbol{s}_4$

which is the structure of Cr_2O_3; and

(3) $\boldsymbol{l}_3 \neq 0,\ \boldsymbol{l}_1 = \boldsymbol{l}_2 = \boldsymbol{m} = 0,\ \boldsymbol{s}_1 = \boldsymbol{s}_2 = -\boldsymbol{s}_3 = -\boldsymbol{s}_4$

In the actual crystal, owing to the relativistic interactions which yield the term $l_{1x}m_y - l_{1y}m_x$, in α-Fe_2O_3 a small vector $\boldsymbol{m}$ arises in addition to the basic antiferromagnetic spin distribution described by the vector $\boldsymbol{l}_1$, and the crystal becomes ferromagnetic. This ferromagnetism will be of relativistic, and not exchange, nature.

For the same reason the antiferromagnetic structure of Cr_2O_3 is determined by the vectors $\boldsymbol{l}_2$ and $\boldsymbol{l}_3$ taken together (for the structure of Cr_2O_3 see below).

* We shall assume that B is always greater than zero, since otherwise our crystal would be the usual exchange ferromagnetic.

FERROMAGNETISM OF α-Fe_2O_3

We shall now consider the magnetic structure of α-Fe_2O_3 in greater detail. Near the transition point $\boldsymbol{l}_1$ and $\boldsymbol{m}$ are small and to calculate them it is sufficient to take into account in the expansion for the thermodynamic potential the second-order terms together with the exchange term proportional to l_1^4 since the remaining fourth-order terms are small compared with the latter. Then Φ can be written as:

$$\Phi = \frac{A}{2}l^2 + \frac{B}{2}m^2 + \frac{\alpha}{2}l_z^2 + \frac{b}{2}m_z^2 + \beta(l_x m_y - l_y m_x) + \frac{C}{4}l^4$$

with the subscript 1 of the vectors $\boldsymbol{l}$ omitted.

The minimum of Φ for a constant $|\boldsymbol{l}|$ is determined by the equations:

$$\begin{aligned} \beta m_y &= \lambda l_x & B m_x - \beta l_y &= 0 \\ -\beta m_x &= \lambda l_y & B m_y + \beta l_x &= 0 \quad l_x^2 + l_y^2 + l_z^2 = l^2 \\ \alpha l_z &= \lambda l_z & (B+b) m_z &= 0. \end{aligned}$$

This group of equations has two solutions:

(I) $\qquad \boldsymbol{m} = 0,\ l_x = l_y = 0;$

(II) $\ l_z = m_z = 0,\ m_x = (\beta/B) l_y,\ m_y = -(\beta/B) l_x.$

In the I state the ion spins are directed along the crystal axis and ferromagnetism is absent. In the II state the spins lie in the Plane (111), and the spontaneous magnetic moment has the magnitude $m = (|\beta|/B)l$ and is perpendicular to $\boldsymbol{l}$. As pointed out earlier $|\beta|/B \cong 10^{-2}$–10^{-5}, therefore $m/l \cong 10^{-2}$–10^{-5} in agreement with the experimental value $m/l = 2 \times 10^{-4}$. In the II state the spins $\boldsymbol{s}_1$, $\boldsymbol{s}_4$ and $\boldsymbol{s}_2$, $\boldsymbol{s}_3$ no longer cancel out but they are rotated with respect to one another by a small angle of the order of 10^{-4} (Fig. 2).

The direction of the vector $\boldsymbol{l}$ in the plane (111) is not determined by the second-order terms in the expansion of Φ. For this we must have an invariant of the lowest order which may be constructed from l_x and l_y. Such an invariant is $(l_x + il_y)^6 + (l_x - il_y)^6$ and depending on its coefficient the vector $\boldsymbol{l}$ will lie either along one of the second-order axes or in one of the planes of symmetry. All these results are in full agreement with the results obtained earlier from considerations of symmetry.

VOL. 4 1958

Of the two possible structures I and II the one which has the lowest thermodynamic potential will actually exist.

In the I state

$$\Phi_{\mathrm{I}} = \frac{A}{2}l^2 + \frac{\alpha}{2}l^2 + \frac{C}{4}l^4$$

and in the II state

$$\Phi_{\mathrm{II}} = \frac{A}{2}l^2 - \frac{\beta^2}{2B}l^2 + \frac{C}{4}l^4$$

Minimizing them with respect to l^2 we find that

$$\Phi_{\mathrm{I}} = -\tfrac{1}{4}\frac{(A+\alpha)^2}{C},$$

$$\Phi_{\mathrm{II}} = -\tfrac{1}{4}\frac{(A-\beta^2/B)^2}{C}, \quad l^2 \cong -\frac{A}{C}.$$

Since $\alpha \sim \beta \sim (v/c)^2$, then neglecting the terms of the order of $(v/c)^4$ we finally obtain

$$\Phi_{\mathrm{I}} \cong -\tfrac{1}{4}\frac{A^2}{C} - \frac{A\alpha}{2C}, \quad \Phi_{\mathrm{II}} = -\tfrac{1}{4}\frac{A^2}{C}.$$

Consequently one or the other state will arise depending on the sign of α. SHULL *et al.*[7] have shown that at $250°\mathrm{K} < T < 950°\mathrm{K}$ the spins lie in the (111) plane, and at $T < 250°\mathrm{K}$ in the [111] direction, and therefore $\alpha > 0$ at $250°\mathrm{K} < T < 950°\mathrm{K}$ (we note that $A < 0$) and $\alpha > 0$ at $T < 250°\mathrm{K}$.

To find the temperature dependence of l and m near the transition point, we shall assume, as usual, that $A = \lambda(T - T_c)$ and therefore

$$l = \sqrt{\left[\frac{\lambda}{C}(T_c - T)\right]}, \quad m = \frac{|\beta|}{B}l = \frac{|\beta|}{B} \times \sqrt{\left[\frac{\lambda}{C}(T_c - T)\right]}.$$

Away from the transition point the value of l is no longer small and its power expansion is in general, not longer valid. It is possible however, to expand in powers of a unit vector $\boldsymbol{\gamma}$ in the direction of $\boldsymbol{l}$ (it can be easily seen that this expansion is equivalent to an expansion in powers of $(v/c)^2$). The expansion in powers of $\boldsymbol{m}$ is still possible because of m's smallness.

We shall write down invariants of fourth order in γ and $\boldsymbol{m}$, which contain $\boldsymbol{m}$ of power not higher than the second (invariants with higher powers of $\boldsymbol{m}$ are negligibly small). As before we should distinguish between the pure exchange invariants, those of mixed nature, and the relativistic invariants. The ratio of the energy corresponding to the mixed invariants to the energy of the exchange ones is proportional to $(v/c)^2$, and the ratio of the energy corresponding to the relativistic invariants to the exchange ones to $(v/c)^4$, and for this reason the relativistic invariants are sometimes called doubly relativistic.

For a crystal with the symmetry of α-Fe_2O_3 there is only one exchange invariant of the fourth order:

$$(\boldsymbol{m}\boldsymbol{\gamma})^2$$

two mixed ones:

$$(\boldsymbol{m}\boldsymbol{\gamma})m_z\gamma_z, \ \gamma_z^2 m^2$$

and five doubly relativistic ones:

$$\gamma_z[(\gamma_x+i\gamma_y)^3-(\gamma_x-i\gamma_y)^3],$$
$$m_z[(\gamma_x+i\gamma_y)^3+(\gamma_x-i\gamma_y)^3]$$
$$\gamma_z^4, \ (\gamma_x m_y-\gamma_y m_x)^2, \ \gamma_z^2(\gamma_x m_y-\gamma_y m_x).$$

For our purpose only the one exchange invariant and the first three relativistic ones are important, since the rest, as we shall find later, are small compared with them. We shall also need one triply relativistic invariant $(\gamma_x+i\gamma_y)^6+(\gamma_x-i\gamma_y)^6$; the ratio of the energy corresponding to this invariant to the exchange energy is proportional to $(v/c)^6$.

Finally the thermodynamic potential can be written as

$$\left.\begin{aligned}\Phi = \frac{a}{2}\gamma_z^2 + \frac{B}{2}m^2 + q(\gamma_x m_y - \gamma_y m_x) + \\ + \frac{b}{2}m_z^2 + \frac{D}{2}(\boldsymbol{\gamma}\boldsymbol{m})^2 + \frac{d}{2i}\gamma_z[(\gamma_z+i\gamma_y)^3 - \\ -(\gamma_x-i\gamma_y)^3] + \frac{f}{2}m_z[(\gamma_x+i\gamma_y)^3+(\gamma_x-i\gamma_y)^3] + \\ + \frac{e}{2}[(\gamma_x+i\gamma_y)^6+(\gamma_x-i\gamma_y)^6] + \frac{g}{4}\gamma_z^4.\end{aligned}\right\} \quad (3)$$

We shall now find the magnetic structure of

α-Fe_2O_3 at temperatures between the point of transition into the paramagnetic state and the point of the I$\rightleftarrows$II state transition, but not too near these points. In this region $a \gg d, e, f, g$, since $d/a \sim f/a \sim g/a \sim (v/c)^2$, $e/a \sim (v/c)^4$.

We shall re-write equation (3) for Φ using polar co-ordinates for the vector $\boldsymbol{\gamma}$ and omitting the term $(g/4)\gamma_z^4$, since the latter is, as shown later, negligible in this region:

$$\left.\begin{aligned}\Phi = &\frac{a}{2}\cos^2\theta + \frac{B}{2}m^2 + \\ &+ q\sin\theta(m_y\cos\phi - m_x\sin\phi) + \\ &+ \frac{b}{2}m_z^2 + \frac{D}{2}[m_z\cos\theta + \\ &+ \sin\theta(m_x\cos\phi + m_y\sin\phi)]^2 + \\ &+ d\cos\theta\sin^3\theta\sin^3\phi + e\cos 6\phi\sin^6\theta + \\ &+ f m_z\sin^3\theta\cos 3\phi.\end{aligned}\right\} \quad (3a)$$

Minimizing Φ with respect to $\boldsymbol{m}$ for given values of θ and ϕ, we find

$$\left.\begin{aligned} m_x &= \frac{q}{B}\sin\theta\sin\phi \\ m_y &= -\frac{q}{B}\sin\theta\cos\phi \\ m_z &= -\frac{f}{B}\sin^3\theta\cos 3\phi.\end{aligned}\right\} \quad (4)$$

At the minimum of Φ we can have the following three sets of values of

$$\left.\begin{aligned} &\text{(I)} && \theta = 0,\ \boldsymbol{m} = 0 \\ &\text{(II)} && \theta \cong \frac{\pi}{2} - \frac{d}{a},\ \phi = \pi/2,\ m_x = \frac{q}{B}, \\ & && m_y = m_z = 0 \\ &\text{(III)} && \theta = \frac{\pi}{2},\ \phi = 0,\ m_y = -\frac{q}{B}, \\ & && m_z = -\frac{f}{B},\ m_x = 0.\end{aligned}\right\} \quad (5)$$

Thus, at not too small values of a, the α-Fe_2O_3 crystal may exist in three magnetic states. In the I state all the spins are directed along the [111] axis and there is no ferromagnetism. In the II state the antiferromagnetic component of the spins lies in one of the planes of symmetry making a small ($\cong d/a \sim (v/c)^2$) angle with the (111) plane, the spontaneous magnetic moment $m_0 = |q|/B \sim (v/c)^2$ is directed along a second-order axis perpendicular to the antiferromagnetic spin component. In the III state the antiferromagnetic component of the spins is directed along one of the second-order axes; the spontaneous magnetic moment is of the same magnitude as in the II state, i.e. $m_0 = |q|/B$, and it lies in the plane of symmetry which is perpendicular to the second-order axis of the antiferromagnetic component making a small angle (of the order of $f/q \sim (v/c)^2$) with the (111) plane.

Once more the results obtained are in full agreement with those found earlier from the considerations of symmetry.

According to NÉEL and PAUTHENET,[1] $m_0 = 0{\cdot}4$ c.g.s. e.m.u.; the reciprocal of the magnetic susceptibility in the paramagnetic phase* is $B = 5\times10^4$. Hence we find that $q = 2\times10^4$ c.g.s. e.m.u.

At a given temperature and pressure of the three states the one with the lowest value of Φ will exist. Using equation (5) we find

$$\Phi_{\mathrm{I}} = \frac{a}{2},\ \Phi_{\mathrm{II}} = -\frac{q^2}{2B} - \frac{d^2}{4a} - e,\ \Phi_{\mathrm{III}} = -\frac{q^2}{2B} + e.$$

Since, in general $|a| \gg q^2/2B, d^2/4|a|, e$, then at $a > 0$ the state II will exist provided $e > -d^2/8a$ (which according to reference 7 occurs at $250°\mathrm{K} < T < 950°\mathrm{K}$), while the state III will exist if $e < -d^2/8a$. When $a < 0$ the state I exists (according to references 1, 7 this occurs at $T < 250°\mathrm{K}$).

At a certain value of a close to zero the transition between the I state and the II (or III) state occurs. The character of this transition can be found from the considerations of symmetry without solving the appropriate equations. From LANDAU's theory it follows that a transition of the second kind is possible only between two states in which the symmetry group of one is the subgroup of the symmetry group of the other. In the section

* Actually in the paramagnetic phase the thermodynamical potential Φ has the form $\Phi = Bm^2/2 + bm_z^2/2$. Therefore, since $\partial\Phi/\partial\boldsymbol{m} = \boldsymbol{H}$, we find $m_x = H_x/B$, $m_y = H_y/B$, $m_z = H_z/(B+b)$ for $b/B \ll 1$.

VOL. 4 1958

on magnetic symmetry we found the symmetry classes of the states I, II and III. It is clear that a transition of the second kind is impossible between the states I and III, and the states II and III. The symmetry class of the II state is a subgroup of the symmetry class of the I state. For a transition of the second kind between the states I and II to be possible absence of third-order terms[4,5] in the expansion of the thermodynamic potential near the transition point is also necessary. In the I to II state transition the x- and y-components of the vector $\boldsymbol{\gamma}$, which occur in the Φ expansion, are small. It can be easily verified that the magnetic symmetry of the I state permits the existence of a third-order invariant $(\gamma_x+i\gamma_y)^3-(\gamma_x-i\gamma_y)^3$, and consequently the I$\rightleftarrows$II transition is also a transition of the first kind.

We shall consider now the behaviour of α-Fe_2O_3 near the point of transition of the first kind from the II to the I state. In this case in the expansion of the thermodynamic potential in equation (3) it is sufficient to retain only the following terms:

$$\Phi=\frac{a}{2}\gamma_z^2+\frac{B}{2}m^2+q(\gamma_x m_y-\gamma_y m_x)+\frac{D}{2}(m\gamma)^2+$$
$$+\frac{d}{2i}[(\gamma_x+i\gamma_y)^3-(\gamma_x-i\gamma_y)^3]\gamma_z+\frac{g}{4}\gamma_z^4.$$

Expressing $\boldsymbol{m}$ in terms of γ and introducing polar co-ordinates for γ, we find

$$\Phi=-\frac{q^2}{2B}+\frac{a'}{2}\cos^2\theta+\frac{g}{4}\cos^4\theta+$$
$$+d\cos\theta\sin^3\theta\sin 3\phi$$

where

$$a'=a+\frac{q^2}{B}.$$

In the I state $\cos\theta=1$ and in the II state $\sin 3\phi=\pm 1$, therefore, for $\theta<\pi/2$, we have

$$\Phi_{\mathrm{I}}=-\frac{q^2}{2B}+\frac{a'}{2}+\frac{g}{4}$$

$$\Phi_{\mathrm{II}}=\Phi_{\mathrm{I}}-\frac{a'}{2}\sin^2\theta-\frac{g}{4}\times$$
$$\times(1+\cos^2\theta)\sin^2\theta-|d|\cos\theta\sin^3\theta.$$

The angle θ is found from the following equation

$$a'\cos\theta+g\cos^3\theta-|d|\sin^3\theta+3|d|\cos^2\theta\sin\theta=0.$$

At the transition point $\Phi_{\mathrm{I}}=\Phi_{\mathrm{II}}$, $\partial\Phi_{\mathrm{II}}/\partial\theta=0$, and these conditions yield the following two equations for determination of the angle θ_c and the quantity a'_c

$$|d|\tan^3\theta_c-a_c'\tan^2\theta_c-3|d|\tan\theta_c-a_c'-g=0,$$

$$\left(a_c'+\frac{g}{2}\right)\tan^2\theta_c+2|d|\tan\theta_c+a_c'+g=0$$

and therefore

$$|d|\tan^2\theta_c+\frac{g}{2}\tan\theta_c-|d|=0$$

and

$$\tan\theta_c=-g/4|d|+\sqrt{(g^2/16d^2+1)}.$$

We shall now calculate the spontaneous magnetic moment m_c at the transition point and the entropy of the transition. Using equation (4) we find

$$m_c=\frac{|q|}{B}\sin\theta_c$$

or

$$\sin\theta_c=\frac{m_c}{m_0}$$

where m_0 is the value of the magnetic moment far from the transition point.

The change of entropy is

$$\Delta S_{\mathrm{I,II}}=-\frac{\partial}{\partial T}(\Phi_{\mathrm{I}}-\Phi_{\mathrm{II}})=-\tfrac{1}{2}\left(\frac{da'}{dT}\right)_c\sin^2\theta_c$$

and the heat of the transition is

$$q_{\mathrm{I,II}}=-\frac{T_c}{2}\left(\frac{da'}{dT}\right)_c\left(\frac{m_c}{m_0}\right)^2.$$

We can now describe fully the behaviour of the α-Fe_2O_3 crystal with change of temperature.

At temperatures $T>950°\mathrm{K}$ the crystal is paramagnetic. At the point $T_c=950°\mathrm{K}$ there is a transition into the state in which the spins of the Fe^{+3} ions differ, as a first approximation, in sign

only: $\boldsymbol{s}_1 = -\boldsymbol{s}_2 = -\boldsymbol{s}_3 = \boldsymbol{s}_4$, and they are directed along the line of intersection of the plane (111) and one of the planes of symmetry. In a higher approximation the spins $\boldsymbol{s}_1$, $\boldsymbol{s}_4$ and $\boldsymbol{s}_2$, $\boldsymbol{s}_3$ are found to be rotated towards one another about the crystal axis [111] by a small angle of the order of 10^{-4}, as shown in Fig. 2, and this produces a spontaneous magnetic moment whose magnitude is 0·04 per cent of the nominal value and which is directed along a second-order axis. Such an explanation for occurrence of ferromagnetism was first put forward by BOROVIK-ROMANOV and ORLOVA.[2] Also the spins do not actually lie in the (111) plane, but make with it a small angle of the order of $(v/c)^2$

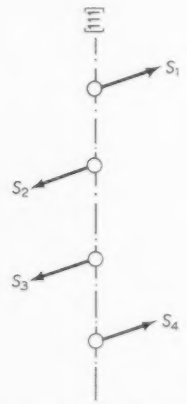


FIG. 3. Projections of the α-Fe_2O_3 spins on a symmetry plane.

(Fig. 3). On approaching $T_c = 250°$K from the direction of high temperatures the angle between the spins and the [111] axis begins to decrease rapidly. At $T_c = 250°$K it reaches a certain critical value θ_c and the spontaneous magnetic moment falls to a value m_c. At $T_c = 250°$K a transition of the first kind occurs when the spins suddenly change their direction to that of [111] and the spontaneous magnetic moment disappears. At $T < 250°$K the crystal becomes antiferromagnetic with the spins oriented along the crystal axis; ferromagnetism is absent in this state.

BEHAVIOUR OF α-Fe_2O_3 IN MAGNETIC FIELDS

Behaviour of the ferromagnetic modification of α-Fe_2O_3 in magnetic fields is of great interest since its magnetic properties in the antiferromagnetic state I are those of a usual antiferromagnetic.

We shall limit ourselves to not too strong fields so that $H^2 \ll aB$, since at fields $H^2 \cong aB$ there may occur a transition from the II into the I state accompanied by destruction of ferromagnetism. For this to occur in known antiferromagnetics at temperatures of the order of hundreds of degrees fields of about 10^5—10^6 oersted are required.

In the presence of an external magnetic field the thermodynamic potential has the form:†

$$\left.\begin{aligned}\tilde{\Phi} = \Phi - \boldsymbol{m}\boldsymbol{H} = {} & \frac{a}{2}\cos^2\theta + \frac{B}{2}m^2 + \\ & + \frac{b}{2}m_z^2 + q\sin\theta(m_y\cos\phi - \\ & - m_x\sin\phi) + \frac{D}{2}[m_z\cos\theta + \\ & + \sin\theta(m_x\cos\phi + m_y\sin\phi)]^2 + \\ & + d\cos\theta\sin^3\theta\sin 3\phi + e\sin^6\theta\cos 6\phi + \\ & + f m_z\sin^3\theta\cos^3\phi - \boldsymbol{m}\boldsymbol{H}.\end{aligned}\right\} \quad (6)$$

We shall first consider the limiting case of very weak fields whose intensity is small compared with the energy of anisotropy in the (111) plane. The latter is given by the terms $d\cos\theta\sin^3\theta\sin 3\phi$ and $e\sin^6\theta\cos 6\phi$. When $H^2 \ll aB$ the value of $\cos\theta$ is still small and of the order of d/a, and therefore the energy of anisotropy is of the order of $e + d^2/a$. The energy of the applied magnetic field is on the whole due to the energy of the spontaneous magnetic moment $m_0 = |q|/B$ and is therefore $H|q|/B$. The criterion of weakness of the applied magnetic field can therefore be expressed as

$$H|q|/B \ll e^2 + d/a \text{ or } H \ll H^* = 6B(e + d^2/4a)/|q|$$

(we shall find useful later the quantity H^*).

In absence of a magnetic field the crystal is in the II state. Its spontaneous magnetic moment is then directed along the x-axis, $\phi = \pi/2$ and

† Φ is the thermodynamic potential in terms of the variables p, T, $\boldsymbol{m}$, while $\tilde{\Phi}$ is this potential in terms of p, T, $\boldsymbol{H}$, i.e. $(d\tilde{\Phi})p$, $T = -\boldsymbol{m}d\boldsymbol{H}$, and it is $\tilde{\Phi}$ that should be minimized at a given value of $\boldsymbol{H}$.

VOL. 4 1958

$\theta = \pi/2 - d/a$. In a weak magnetic field the thermodynamic potential $\tilde{\Phi}$ is a quadratic function of the field intensity components, and the quantities $m_x - |q|/B$, m_y, m_z, $\pi/2-\phi$, $\pi/2 - d/a - \theta$ are linear functions of the field intensity. Introducing the quantities

$$m_x' = m_x - |q|/B,\ \kappa = \pi/2 - \phi,\ \alpha = \pi/2 - d/a - \theta$$

we expand $\tilde{\Phi}$ neglecting powers higher than the squares of these quantities:

$$\tilde{\Phi} = -\frac{q^2}{2B} - \frac{d^2}{2a} - e + \frac{a}{2}\alpha^2 + \frac{dq}{a}\alpha m_x' +$$
$$+\left(\frac{q^2}{2B} + \frac{Dq^2}{2B^2} + \frac{9d^2}{2a} + 18e\right)\kappa^2 + q\left(1 + \frac{D}{B}\right)\kappa m_y +$$
$$+\left(\frac{Dqd}{Ba} - 3f\right)\kappa m_z + \frac{B}{2}m_x'^2 + \frac{B+D}{2}m_y^2 +$$
$$+\frac{B+b}{2}m_z^2 + \frac{Dd}{a}m_y m_z - H_x m_x' - H_y m_y - H_z m_z.$$

We also introduce $m_y' = m_y - (q/B)\kappa$ in place of m_y. Then $\tilde{\Phi}$ becomes:

$$\tilde{\Phi} = \frac{a}{2}\alpha^2 + \frac{dq}{a}\alpha m_x' + \frac{B}{2}m_x' - H_x m_x' + \frac{B+D}{2}m_y'^2 +$$
$$+\frac{B+b}{2}m_z^2 + \frac{Dd}{a}m_z m_y' - 3f\kappa m_z + \frac{q}{B}H_y\kappa +$$
$$+18e'\kappa^2 - H_y m_y' - H_z m_z,$$

where $e' = e + d^2/4a$ and the constant terms are omitted.

Determining α, κ, m_x', m_y', m_z from the condition of minimum of Φ we obtain the following system of equation:

$$a\alpha + dqm_x'/a = 0,$$
$$Bm_x' + dq\alpha/a - H_x = 0,$$
$$36e'\kappa + qH_y/B - 3fm_z = 0,$$
$$(B+D)m_y' + Ddm_z/a - H_y = 0,$$
$$(B+b)m_z + Ddm_y'/a - 3f\kappa - H_z = 0,$$

from which by a simple calculation, retaining only the largest terms, we find for κ and α:

$$\left.\begin{aligned}\kappa &= -qH_y/36e'B + fH_z/12e'B,\\ \alpha &= -dqH_x/a^2B,\end{aligned}\right\} \quad (7)$$

for the components of the magnetic moment:

$$\left.\begin{aligned}m_x &= \frac{|q|}{B} + \frac{H_x}{B},\\ m_y &= \frac{q^2H_y}{36e'B^2} - \frac{qf}{12e'B^2}H_z,\\ m_z &= \frac{H_z}{B+b} - \frac{qf}{12e'B^2}H_y,\end{aligned}\right\} \quad (8)$$

and for the magnetic susceptibility:

$$\chi_{xx} = \frac{1}{B},\ \chi_{yy} = \frac{q^2}{36e'B^2},\ \chi_{yz}$$
$$= -\frac{qf}{12e'B^2},\ \chi_{zz} = \frac{1}{B+b},\ \chi_{xy} = \chi_{xz} = 0.$$

χ_{xx}, χ_{zz}, χ_{yz} are of the same order as the susceptibility, $1/B$ in the paramagnetic phase, but χ_{yy} has a much larger value of the order $(v/c)^{-2}$.

It is not by accident that χ_{xy} and χ_{xz} are equal to zero. The thermodynamic potential $\tilde{\Phi}$, as a function of the field intensity components, should remain invariant under all transformations of the crystal symmetry. We found earlier that the symmetry group of the II state consists of the elements U_2, I, σ_d. The most general combination of H_x, H_y, H_z invariant under these three transformations is

$$\tilde{\Phi} = -m_0 H_x - \tfrac{1}{2}\chi_{xx}H_x^2 - \tfrac{1}{2}\chi_{yy}H_y^2 -$$
$$-\tfrac{1}{2}\chi_{zz}H_z^2 - \chi_{yz}H_y H_z$$

from which we obtain for the magnetic moment an expression of the type given in equation (8).

We shall consider the other limiting case of fields $H \gg H^*$. In this case the terms representing the anisotropy energy in the expression for the thermodynamic potential can be neglected. Also, since we are assuming that $H^2 \ll aB$, we can neglect the small departure of the angle θ from the right angle. Now $\tilde{\Phi}$ has the following form:

$$\tilde{\Phi} = \frac{B}{2}m^2 + \frac{b}{2}m_z^2 + q(m_y\cos\phi - m_x\sin\phi) +$$
$$+\frac{D}{2}(m_x\cos\phi + m_y\sin\phi)^2 - \boldsymbol{mH}.$$

We shall express the magnetic moment $\boldsymbol{m}$ in terms of the external field and the angle ϕ:

$$m_x = \frac{q}{B}\sin\phi + \frac{H_x}{B} - \frac{D}{B(B+D)} \times$$
$$\times (H_x \cos\phi + H_y \sin\phi)\cos\phi,$$

$$m_y = -\frac{q}{B}\cos\phi + \frac{H_y}{B} - \frac{D}{B(B+D)} \times$$
$$\times (H_x \cos\phi + H_y \sin\phi)\sin\phi,$$

$$m_z = \frac{H_z}{B+b}.$$

In the plane xy we introduce polar co-ordinates for H_x and H_y: $H_x = H_t\cos\psi$ and $H_y = H_t\sin\psi$ and resolve $\boldsymbol{m}$ in the direction of $\boldsymbol{H}_t$ and at right angles to the latter:

$$\left.\begin{aligned} m_{\parallel} &= -\frac{q}{B}\sin(\psi-\phi) + \frac{H_t}{B} - \frac{DH_t}{B(B+D)} \times \\ &\quad \times \cos^2(\psi-\phi) \\ m_{\perp} &= \frac{q}{B}\cos(\psi-\phi) - \frac{DH_t}{B(B+D)} \times \\ &\quad \times \cos(\psi-\phi)\sin(\psi-\phi) \end{aligned}\right\} \quad (9)$$

Excluding $\boldsymbol{m}$ from the expression for the thermodynamic potential and omitting the terms not dependent on ϕ, we find:

$$\tilde{\Phi} = \frac{q}{B}H_t\sin(\psi-\phi) + \frac{DH_t^2}{2B(B+D)}\cos^2(\psi-\phi). \quad (10)$$

The angle ϕ as a function of $\boldsymbol{H}_t$ is found from the following equation:

$$\frac{\partial\tilde{\Phi}}{\partial\phi} = -\frac{q}{B}H_t\cos(\psi-\phi) +$$
$$+\frac{DH_t^2}{B(B+D)}\cos(\psi-\phi)\sin(\psi-\phi) = 0,$$

which has two solutions:

(1) $$\cos(\psi-\phi) = 0$$

(2) $$\sin(\psi-\phi) = \frac{q(B+D)}{DH_t}.$$

The following values of the thermodynamic potential correspond to the above two solutions:

$$\tilde{\Phi}_1 = -\frac{|q|}{B}H_t, \quad \tilde{\Phi}_2 = \frac{DH_t^2}{2B(B+D)} + \frac{q^2(B+D)}{2BD},$$

and the sign of the difference of the above two values

$$\tilde{\Phi}_2 - \Phi_2 = \frac{D}{2B(B+D)}\left[H_t + \frac{|q|(B+D)}{D}\right]^2$$

is determined by the sign of D.

Thus in fields $H \gg H^*$ the crystal may exist in two states depending on the sign of D. If $D > 0$, then

$$\phi = \psi + \frac{\pi}{2}$$

and

$$m_{\parallel} = |q|/B + H_t/B, \quad m_{\perp} = 0, \quad m_z = H_z/(B+b).$$

The spontaneous magnetic moment which lies in the xy-plane is oriented along the field and the antiferromagnetic components of the spins perpendicularly to the field. In contrast to the usual antiferromagnetics the susceptibility is now practically isotropic. This was noted for α-Fe_2O_3 by NÉEL and PAUTHENET[1] and therefore favours the positive sign of D.

For $D < 0$ we find from equation (9) that

$$m_{\parallel} = H_t/(B+D), \quad m_{\perp} = 0, \quad m_z = H_z/(B+b).$$

The susceptibility is now strongly anisotropic in the direction of the crystal axis and there is no spontaneous magnetic moment, i.e. the crystal becomes a usual antiferromagnetic. Figs. 4(a) and 4(b) show the dependence of $m_{\parallel}$ on the field for various signs of D.

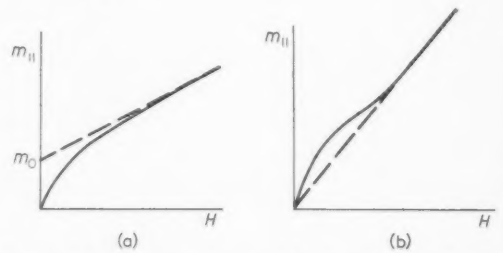


FIG. 4. Magnetization curves: (a)—$D > 0$. (b)—$D < 0$.

VOL. 4 1958

We shall finally consider the region of fields when $H \cong H^*$. In this case simple but tedious calculations show that we must add an anisotropy energy term $e' \cos 6\phi$ to the thermodynamic potential given by equation (10):

$$\tilde{\Phi} = \frac{q}{B} H_t \sin(\psi - \phi) + \frac{DH_t^2}{2B(B+D)} \cos^2(\psi - \phi) + e' \cos 6\phi. \tag{11}$$

The magnetic moment is still given by equation (9).

For $H_t \gtrsim H^*$ the second term in equation (11) is small compared with the first, since $H^* = 6Be'/|q|$ and $qH^*/B \sim (v/c)^6$, while $H^{*2}/B \sim (v/c)^8$. Thus $\tilde{\Phi}$ can now be written as:

$$\tilde{\Phi} = \frac{q}{B} H_t \sin(\psi - \phi) + e' \cos 6\phi \tag{12}$$

and the equation, which determines ϕ, as:

$$qH_t \cos(\psi - \phi) + 6e' \sin 6\phi = 0.$$

The above equation and equation (12) may be reduced to a non-dimensional form on replacing H_t by $h = H_t/H^*$

$$\left.\begin{aligned} \tilde{\Phi}/e' &= 6h \sin(\psi - \phi) + \cos 6\phi \\ h \cos(\psi - \phi) &+ \sin 6\phi = 0. \end{aligned}\right\} \tag{13}$$

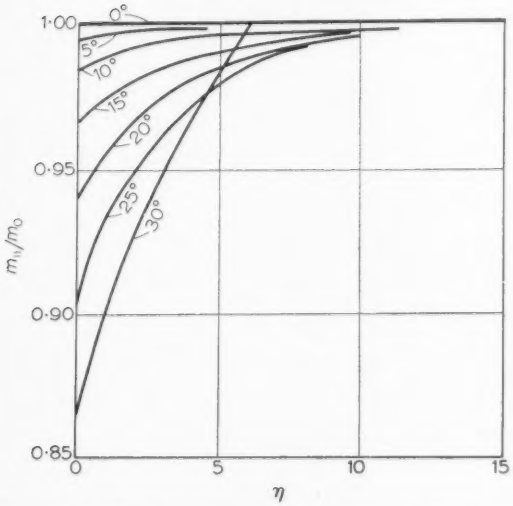


FIG. 5. Equilibrium magnetization curves for various external field directions. Figures denote value of ψ.

Fig. 5 shows equilibrium curves of magnetization calculated by numerical solution of equations (13). The ordinate axis gives values of $m_{\parallel} - H_t/B$ in units of the spontaneous magnetic moment $m_0 = |q|/B$.

In addition to the equilibrium solutions equations (13) contain also non-equilibrium solutions of the hysteresis type. The region outside which hysteresis is always impossible is given by the two conditions $\partial\tilde{\Phi}/\partial\phi = \partial^2\Phi/\partial\phi^2 = 0$ which yield the equations of the bounding curve in a parametric form

$$\left.\begin{aligned} h \cos(\psi - \phi) &+ \sin 6\phi = 0 \\ h \sin(\psi - \phi) &+ 6 \cos 6\phi = 0. \end{aligned}\right\} \tag{14}$$

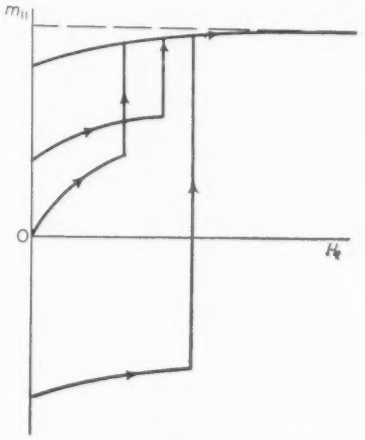


FIG. 6. Typical hysteresis curves.

Fig. 6 shows typical hysteresis curves given by equations (13) and the conditions (14).

Cr_2O_3

We shall briefly consider the antiferromagnetic structure of Cr_2O_3. The spin distribution for the Cr^{3+} ions is described by the vectors $\boldsymbol{l}_2$ and $\boldsymbol{l}_3$. It was shown earlier that the components of $\boldsymbol{l}_2$ and $\boldsymbol{l}_3$ transform in the crystal symmetry transformations in the same way as those of $\boldsymbol{l}_1$ and $\boldsymbol{m}$, which describe the magnetic structure of α-Fe_2O_3, except that $\boldsymbol{l}_2$ and $\boldsymbol{l}_3$ change their sign in inversion transformations while $\boldsymbol{l}_1$ and $\boldsymbol{m}$ do not. Since in the thermodynamic potential expansion $\boldsymbol{l}_2$ and $\boldsymbol{l}_3$ occur only in the even powers their inversion properties

will not affect the expansion and Φ for Cr_2O_3 may be obtained from Φ for α-Fe_2O_3 by replacing $\boldsymbol{l}_1$ by $\boldsymbol{l}_2$ and $\boldsymbol{m}$ by $\boldsymbol{l}_3$. Without repeating the calculations we can now describe the thermodynamically possible antiferromagnetic structures of Cr_2O_3.

As a first approximation the structure of Cr_2O_3 may be described by the vector $\boldsymbol{l}_2$. All the four Cr^{3+} ions possess spins of the same magnitude and direction, and $\boldsymbol{s}_1 = -\boldsymbol{s}_2 = \boldsymbol{s}_3 = -\boldsymbol{s}_4$ (Fig. 1c). As before, three states are possible which differ by the orientation of the spins $\boldsymbol{s}$ with respect to the crystal axes. In the I state the spins are directed along the [111] crystal axis, in the II state the spins lie along the line of intersection of one of the symmetry planes with the (111) plane, and in the III state the spins are directed along one of the second-order axes.

In a higher approximation the spins $\boldsymbol{s}_1$, $\boldsymbol{s}_2$ and the spins $\boldsymbol{s}_3$, $\boldsymbol{s}_4$ in the II and the III states rotate about the [111] axis towards one another by a small angle as shown in Fig. 7. Their sum still remains

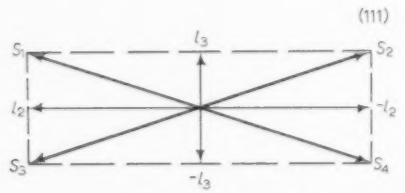


FIG. 7. Projections of the Cr_2O_3 spins on (111).

equal to zero so that no ferromagnetism arises here [more exactly $\boldsymbol{s}_1 = \frac{1}{4}(\boldsymbol{l}_2+\boldsymbol{l}_3)$, $\boldsymbol{s}_2 = \frac{1}{4}(-\boldsymbol{l}_2+\boldsymbol{l}_3)$, $\boldsymbol{s}_3 = \frac{1}{4}(\boldsymbol{l}_2-\boldsymbol{l}_3)$, $\boldsymbol{s}_4 = -\frac{1}{4}(\boldsymbol{l}_2+\boldsymbol{l}_3)$].

The magnetic properties of Cr_2O_3 will not differ from those of usual antiferromagnetics and we shall not discuss them.

$MnCO_3$, $CoCO_3$, $FeCO_3$

In the isomorphous antiferromagnetic crystals $MnCO_3$, $CoCO_3$ and $FeCO_3$ the weak ferromagnetism is exhibited by $MnCO_3$ and $CoCO_3$ but not by $FeCO_3$ which is a pure antiferromagnetic.

Ferromagnetism of $MnCO_3$ and $CoCO_3$ was studied in detail by BOROVIK-ROMANOV and ORLOVA.[2] According to their measurements the spontaneous magnetic moment of $MnCO_3$ is 0·2 per cent of the nominal value and that of $CoCO_3$, 2 per cent. $FeCO_3$ was studied by BIZETTE[12] who found that in the antiferromagnetic state the spins of the Fe^{+2} ions are directed along the [111] crystal axis.

The properties of these crystals can be explained if we assume that the antiferromagnetic transition occurs in them without a change of the unit cell. $MnCO_3$, $CoCO_3$ and $FeCO_3$ belong to the rhombohedral system and their symmetry is given by the space group D_{3d}^6,[6] i.e. by the same group as α-Fe_2O_3 and Cr_2O_3. The two metal ions in the unit cell lie on the volume diagonal of the rhombohedron at the points (1/4, 1/4, 1/4) and (3/4, 3/4, 3/4).

As before we shall describe the magnetic structure in terms of the ion spins $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$. We shall introduce vectors $\boldsymbol{m} = \boldsymbol{s}_1+\boldsymbol{s}_2$ and $\boldsymbol{l} = \boldsymbol{s}_1-\boldsymbol{s}_2$. It can be easily shown that under all the symmetry transformations $\boldsymbol{m}$ and $\boldsymbol{l}$ transform independently of one another, $\boldsymbol{m}_z$ and $\boldsymbol{l}_z$ transform according to unidimensional representations A_{2g} and A_{1g} of the point group D_{3d} and their x- and y-components according to one representation E_g, that is the vectors $\boldsymbol{m}$ and $\boldsymbol{l}$ here transform in the same way as the vectors $\boldsymbol{m}$ and $\boldsymbol{l}_1$ of α-Fe_2O_3.

It follows therefore that the magnetic properties of the crystals considered in this section are identical with those of α-Fe_2O_3 and therefore all the results obtained earlier apply here.

In particular the fact that $FeCO_3$ does not exhibit ferromagnetism is fully explained by the direction of the Fe^{2+} ions spins along the crystal axis (the I state).

From the occurrence of ferromagnetism in $MnCO_3$ and $CoCO_3$ we can conclude that they exist in the states (II or III) with their spins in the (111) plane.

All calculations of the magnetic properties of α-Fe_2O_3 are fully applicable for the ferromagnetic $MnCO_3$ and $CoCO_3$.

Acknowledgements—The author is grateful to Acad. LANDAU and LIFSHITZ for their valuable advice and criticism of the results. The author is also indebted to BOROVIK-ROMANOV for useful discussions.

REFERENCES

1. NÉEL L. and PAUTHENET R. *C.R. Acad. Sci., Paris* **234,** 2172 (1952).
2. BOROVIK-ROMANOV A. S. and ORLOVA M. P. *Zh. Eksp. Teor. Fiz.* **31,** 579 (1956).

VOL. 4 1958

3. YIN-YUAN LI *Phys. Rev.* **101,** 1450 (1956).
4. LANDAU L. D. *Phys. Z. Sowjet.* **11,** 26, 545 (1937).
5. LANDAU L. D. and LIFSHITZ E. M. *Statistical Physics* (in Russian) Moscow (1951).
6. WYCKOFF R. W. G. *The Structure of Crystals.* New York (1931).
7. SHULL C. G., STRAUSER W. A. and WOLLAN E. O. *Phys. Rev.* **83,** 333 (1951).
8. BROCKHOUSE B. N. *J. Chem. Phys.* **21,** 961 (1953).
9. TAVGER B. A. and ZAITSEV V. M. *Z. Exp. Theor. Phys.* **30,** 564 (1956).
10. *Internationalle Tabellen zur Bestimmung von Kristallstrukturen* Vol. 1. (1935).
11. ROSENTHAL J. and MURPHY G. *Rev. Mod. Phys.* **8,** 317 (1936).
12. BIZETTE H. *J. Phys. Radium* No. **12,** 161 (1951).

J. Phys. Chem. Solids. Pergamon Press 1958. Vol. 4. pp. 256–270.

A FERROMAGNETIC DYNAMICAL EQUATION*

HERBERT B. CALLEN

Department of Physics, University of Pennsylvania, Philadelphia, Pa.

(*Received* 7 *August* 1957)

Abstract—Based on general considerations concerning the dissipative mechanism, a ferromagnetic dynamical equation is proposed. This equation contains three parameters, which have definite quantum-mechanical significance. A simple theory for one such parameter is given, and an explicit rotational dynamical equation is obtained in the special case in which the $(k = 0) \rightarrow (k \neq 0)$ spin-wave transitions dominate the dissipative process. It is suggested that certain observed size and shape effects may be explained qualitatively on the basis of this equation.

VOL.
4
1958

1. INTRODUCTION

A FERROMAGNETIC dynamical equation is the magnetic analogue of Ohm's Law; it seeks to describe the response of a ferromagnetic material to an applied field. However Ohm's Law is essentially determined in form by the condition of linearity alone, whereas the three-dimensional character of the magnetic response permits the formulation of many different linear laws. Therefore, in the magnetic case, it is necessary to appeal to specific models to determine the proper form of the dynamical equation. Fortunately the underlying models of ferromagnetic dissipation have been studied extensively.[1,2] As we shall see, at least some aspects of the form of the dynamical equation depend only upon certain very general characteristics of the models, and therefore can be determined with some certainty. By a general examination of various ferromagnetic dissipative models, we shall be led to the formulation of a specific form of dynamical equation.

The importance of determining the form of the dynamical equation stems from both fundamental and practical considerations. From the fundamental point of view, it specifies the basic parameters which are to be measured and for which detailed theories are to be sought; practically, it provides the basis for the phenomenological theories of ferromagnetic resonance, of "switching", and of domains and domain walls.

* Supported by the Office of Naval Research.

A convenient analytic statement of the problem is obtained by expanding the rate of change of the magnetization $\boldsymbol{M}$ in the three orthogonal vectors $\boldsymbol{M}$, $(\boldsymbol{M}\times\boldsymbol{H})$, and $\boldsymbol{M}\times(\boldsymbol{M}\times\boldsymbol{H})$.

$$\dot{\boldsymbol{M}} = \alpha\boldsymbol{M} - \gamma\boldsymbol{M}\times\boldsymbol{H} - \lambda\boldsymbol{M}\times(\boldsymbol{M}\times\boldsymbol{H}). \qquad (1)$$

The three functions α, γ, and λ are unknown scalar functions of $\boldsymbol{M}$, $\boldsymbol{H}$, and the intrinsic characteristics of the sample. The problem is to evaluate the functional dependence of α, γ, and λ on $\boldsymbol{M}$, $\boldsymbol{H}$, and the relevant other variables.

A currently popular ferromagnetic dynamical equation is the LANDAU–LIFSHITZ equation.[3] LANDAU and LIFSHITZ assumed first that the ferromagnetic response consists of a pure rotation, implying that the function α vanishes. In the absence of any further theoretical information, they then merely assumed that the remaining two functions, γ and λ, are constants. The LANDAU–LIFSHITZ equation is, then

$$\dot{\boldsymbol{M}} = -\gamma_L\boldsymbol{M}\times\boldsymbol{H} - \lambda_L\boldsymbol{M}\times(\boldsymbol{M}\times\boldsymbol{H}). \qquad (2)$$

A slightly different way of writing this equation has been suggested by GILBERT.[4] The GILBERT equation is equivalent to the LANDAU–LIFSHITZ equation in that it also takes γ and λ to be constants. However, these constants are written in terms of two other constants γ_G and λ_G, which are alleged to be more fundamental. The Gilbert equation is

$$\dot{\boldsymbol{M}} = -\frac{\gamma_G}{1+\lambda_G^2 M^2}\boldsymbol{M}\times\boldsymbol{H} -$$

$$-\frac{\lambda_G\gamma_G}{1+\lambda_G{}^2M^2}\boldsymbol{M}\times(\boldsymbol{M}\times\boldsymbol{H}). \tag{3}$$

The reason for GILBERT's preference for γ_G and λ_G as fundamental constants lies in the possibility of rewriting equation (3) in the simple form

$$\dot{\boldsymbol{M}}=-\gamma_G\boldsymbol{M}\times\left[\boldsymbol{H}-\frac{\lambda_G}{\gamma_G}\dot{\boldsymbol{M}}\right]. \tag{4}$$

This equation is then subject to the attractive heuristic interpretation that the dissipation introduces an effective magnetic field opposite in direction and proportional in magnitude to $\dot{\boldsymbol{M}}$.

Another current class of ferromagnetic dynamical equations consists of adaptations of the BLOCH equation, originally derived for paramagnetics.[5] These equations typically have two relaxation times; one for the component of the magnetization parallel to the applied field and one for the perpendicular component. The "BLOCH–BLOEMBERGEN equation" is

$$\dot{\boldsymbol{M}}=-\gamma_0\boldsymbol{M}\times\boldsymbol{H}-\frac{\boldsymbol{M}\cdot\boldsymbol{H}_0-M_0H_0}{T_1}\cdot\frac{\boldsymbol{H}_0}{H_0{}^2}-$$

$$-\frac{1}{T_2H_0}(\boldsymbol{M}\times\boldsymbol{H}_0), \tag{5}$$

where T_1 and T_2 are the longitudinal and transverse relaxation times, and $\boldsymbol{H}_0$ is the externally applied field.

The general equation which we first obtain is superficially similar to these equations. But in place of the LANDAU–LIFSHITZ parameter, λ_L, or the two relaxation times, T_1 and T_2, it contains three parameters $\lambda_{0\sigma}$, λ_{0k}, and $\lambda_{k\sigma}$. These parameters are defined in terms of definite quantum-mechanical transition probabilities, of creation and destruction of magnons. The essential point is that these parameters are themselves functions of the applied field, of the state of the system, and of the intrinsic properties of the material. Actual application of the equation requires that these dependences be calculated explicitly. Accordingly we consider the special case in which a single parameter dominates the behavior of the equation, and we give a simple theory of the field- and shape-dependence of this parameter. The resulting explicit dynamical equation may be applicable to the rotational response in disordered ferromagnetic alloys.

2. THE SPIN-WAVE REPRESENTATION

The various models of ferromagnetic dissipative processes, to which we shall make reference, employ the spin-wave representation. We shall briefly review the relevant features of that representation.

The spin-wave description of a ferromagnetic material was introduced by BLOCH[6] on the basis of a HEITLER-LONDON localized orbital model. HERRING and KITTEL[7] have shown how the theory can be reformulated on a phenomenological basis which permits the convenient introduction of anisotropy and demagnetization contributions. More recently, ANDERSON and SUHL[8,2,12] have investigated the demagnetization contribution and have found an important alteration in the spin-wave spectrum. Finally, the validity of the entire spin-wave approximation has been investigated by DYSON,[9] culminating in the happy result of demonstrating that the spin-wave description is more satisfactory than had previously been anticipated.

The essential results of the spin-wave analysis are the following. The complete magnetization or spin field can be characterized by the set of its Fourier amplitudes. Each Fourier component constitutes a spin-wave. For the low-lying states the Hamiltonian is approximately diagonal in the spin-wave co-ordinates and, in fact, is approximately a sum of terms of harmonic oscillator form. The wave function is a product of simple-harmonic oscillator functions for each spin-wave, and the energy is a sum of terms of the usual form

$$E=\sum_{\boldsymbol{k}}(n_{\boldsymbol{k}}+\tfrac{1}{2})\hbar\omega(\boldsymbol{k}). \tag{6}$$

The integers $n_{\boldsymbol{k}}$ are the quantum numbers of the spin-waves, or equivalently stated, $n_{\boldsymbol{k}}$ is the "number of magnons with wave-vector $\boldsymbol{k}$". The wave-vectors $\boldsymbol{k}$ are restricted to the reduced, rationalized, reciprocal lattice in the standard fashion.

The dependence of $\omega(\boldsymbol{k})$ on $\boldsymbol{k}$ has been derived by ANDERSON and SUHL[8,12] for the special case of the applied field parallel to a principal axis of the ellipsoidal sample. In an appendix we show how the general case, in which the applied field has an

arbitrary direction, can be reduced to their special case with a particular effective field applied along the principal axis of an effective ellipsoid.

Let $\boldsymbol{H}_0$ be the applied field and let the sample have a demagnetization tensor $4\pi\mathbf{N}$. We first define a unit vector $\boldsymbol{\zeta}$ in the direction of the equilibrium magnetization. As shown in the appendix, $\boldsymbol{\zeta}$ is determined by the condition that

$$H_0 - 4\pi M_0 \boldsymbol{\zeta} \cdot \mathbf{N} \text{ is parallel to } \boldsymbol{\zeta}. \tag{7}$$

It is then shown in the appendix that the spin-wave spectrum in the actual sample is identical to the spectrum in a hypothetical sample with a demagnetizing tensor $4\pi\mathbb{N}$, where

$$\mathbb{N} = \mathbf{N} - \frac{1}{4\pi M_0}(H_0\boldsymbol{\zeta} + \boldsymbol{\zeta}H_0) + \frac{1}{2\pi M_0}(\boldsymbol{H}_0 \cdot \boldsymbol{\zeta})(\boldsymbol{\zeta}\,\boldsymbol{\zeta}). \tag{8}$$

The direction $\boldsymbol{\zeta}$ is a principal axis of this hypothetical demagnetizing tensor, and an effective field of magnitude $\boldsymbol{H}_0 \cdot \boldsymbol{\zeta}$ is considered to be applied along this axis. The relationship of the associated tensor $\mathbb{N}$ to the original tensor $\mathbf{N}$ is indicated in Fig. 1.

The spin-wave spectrum, for k large compared to the reciprocal of the smallest sample dimension, is then

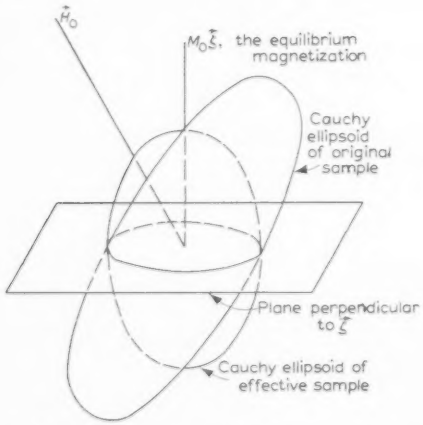


FIG. 1. Actual and "associated" demagnetizing tensors. The Cauchy ellipsoid of a tensor $\mathbf{N}$ is the locus of $\boldsymbol{r}$ if $\boldsymbol{r} \cdot \mathbf{N} \cdot \boldsymbol{r} = 1$. The Cauchy ellipsoids of $\mathbf{N}$ and $\mathbb{N}$ have the same $\boldsymbol{\zeta}$-intercept and the same trace in the plane perpendicular to $\boldsymbol{\zeta}$.

$$\omega^2(\boldsymbol{k}) = \omega_M{}^2(h - N_\zeta + \kappa^2)(h - N_\zeta + \kappa^2 + \sin^2\theta), \tag{9}$$

whereas, for $\boldsymbol{k} = 0$ the natural frequency is

$$\omega^2(0) = \omega_M{}^2(h - N_\zeta + N_\xi)(h - N_\zeta + N_\eta). \tag{10}$$

In these equations we have used the following notation:

$$\omega_M = 4\pi M_0 \gamma_0 \tag{11}$$

$$h = (\boldsymbol{H}_0 \cdot \boldsymbol{\zeta})/4\pi M_0 \tag{12}$$

$$N_\zeta = \boldsymbol{\zeta} \cdot \boldsymbol{N} \cdot \boldsymbol{\zeta}; \tag{13}$$

one of the *eigen*values of $\mathbb{N}$. The other two *eigen*values are N_ξ and N_η.

θ = angle between $\boldsymbol{k}$ and $\boldsymbol{\zeta}$.

$$\boldsymbol{\varkappa} = (JZS_0a^2/12\pi)\boldsymbol{k}. \tag{14}$$

J = exchange integral, Z = number of nearest neighbors of a given spin, S_0 = spin quantum number per ion, and a = inter-ionic distance.

The natural frequencies of those modes for which $\boldsymbol{k}$ is neither zero nor large with respect to the reciprocal sample dimensions have been studied by WALKER.[10] There are sufficiently few such modes so that we will not have to consider them explicitly.

The spin-wave spectrum is shown schematically in Fig. 2.

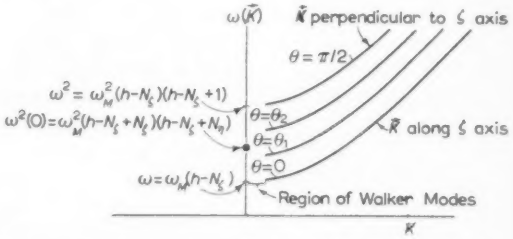


FIG. 2. Spin-wave spectrum in an ellipsoidal sample.

The semi-classical picture[11] of a spin-wave in an infinite medium is indicated in Fig. 3. The spin-wave of wave-vector $\boldsymbol{k}$, occupied by a single magnon, can be pictured as a state in which all spins are inclined by a slight constant angle with respect to the $\boldsymbol{\zeta}$-axis, and precess around the $\boldsymbol{\zeta}$-axis. The total change in the $\boldsymbol{\zeta}$-component of spin

VOL. 4 1958

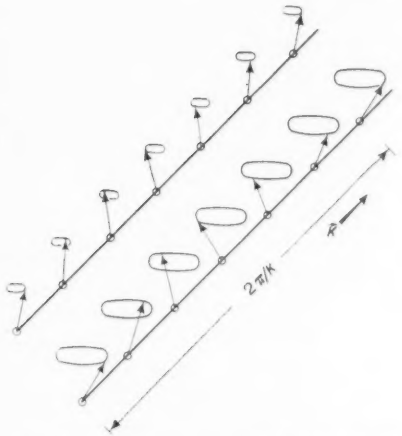


FIG. 3. Semiclassical representation of a spin-wave. The spin-wave is shown occupied by a single magnon and occupied by two magnons.

corresponds to the inversion of a single spin, and the precessional phase varies sinusoidally in space, with a wavelength of $2\pi/k$. If the spin-wave is occupied by two magnons, the angle of each spin is larger, so that the change in ζ-component of spin corresponds to the inversion of two spins.

The spin-wave of zero wave-vector describes the homogeneous precession of all spins, in phase. It is with this mode that any transverse applied magnetic field (such as the r.f. field in a microwave-resonance experiment) will couple.

A theorem of prime importance to our discussion states that each magnon of non-zero $\boldsymbol{k}$ decreases the magnitude of the macroscopic magnetization. That is

$$\langle n_0 \dots n_{\boldsymbol{k}} \dots | \boldsymbol{M} \cdot \boldsymbol{M} | n_0 \dots n_{\boldsymbol{k}} \dots \rangle = M(M+\gamma_0\hbar) \simeq M^2, \tag{15}$$

where

$$M = M_0 - \gamma_0\hbar \sum_{\boldsymbol{k}\neq 0} n_{\boldsymbol{k}} = \boldsymbol{M}_0 - n'\gamma\hbar. \tag{16}$$

Here M_0 is the absolute saturation value of the magnetization,

$\gamma_0\hbar S_0$ is the magnetic moment per ion,

$$n' \equiv \sum_{\boldsymbol{k}\neq 0} n_{\boldsymbol{k}}$$

is the number of magnons with $\boldsymbol{k} \neq 0$, (17)

and

$\langle n_0 \dots n_{\boldsymbol{k}} \dots |$ is the wave-function with n_0 magnons of $\boldsymbol{k} = 0$, $n_{\boldsymbol{k}}$ magnons of wave-vector $\boldsymbol{k}$, etc.

Equations (15) and (16) relate the number of non-zero wave-vector magnons to the macroscopically observable magnetization.

The number of zero-wave-vector magnons can be related to a macroscopic observable also. In an infinite medium each magnon excited reduces the ζ-component of the spin by unity, whence

$$\langle n_0 \dots n_{\boldsymbol{k}} \dots | M_\zeta | n_0 \dots n_{\boldsymbol{k}} \dots \rangle = M_0 - \gamma_0\hbar \sum_{\boldsymbol{k}} n_{\boldsymbol{k}} = M_0 - (n_0 + n')\gamma_0\hbar \tag{18}$$

or, by equation (16),

$$M_\zeta = M - n_0\gamma_0\hbar. \tag{19}$$

In a finite ellipsoidal sample, equations (15) and (16) remain valid, but equations (18) and (19) are not correct. The reason is that the $\boldsymbol{k} = 0$ spin-wave senses the transverse demagnetization coefficients, and precesses in an elliptical rather than a circular cone. The analogue of equation (18) consequently must take account of the transverse demagnetization coefficients. The magnons of $\boldsymbol{k} \neq 0$, however, do not sense the transverse demagnetization coefficients (compare equation (10) with equation (9)) and therefore equation (16) remains valid.

The simplest way of obtaining the correct form of equations (18) and (19) is to appeal to energy considerations. Consider an ellipsoidal sample in which the magnetization has been rotated homogeneously away from its equilibrium direction. The increase in energy associated with the rotation is

$$E - E_0 = [-\boldsymbol{M} \cdot \boldsymbol{H}_0 + 2\pi \boldsymbol{M} \cdot \mathbf{N} \cdot \boldsymbol{M}] - [-M\boldsymbol{\zeta} \cdot \boldsymbol{H}_0 + 2\pi M^2 \boldsymbol{\zeta} \cdot \mathbf{N} \cdot \boldsymbol{\zeta}]. \tag{20}$$

To second order in the small deviation $\boldsymbol{m}_0$ we write

$$\boldsymbol{M} = M\left(1 - \frac{m_0^2}{2M^2}\right)\boldsymbol{\zeta} + \boldsymbol{m}_0 \tag{21}$$

where $\boldsymbol{m}_0$ is perpendicular to $\boldsymbol{\zeta}$.

Substituting this value for $\boldsymbol{M}$, we find, by virtue of equation (7), that to second order in $\boldsymbol{m}_0$

$$E - E_0 = 2\pi m_0^2(h - N_\zeta) + 2\pi \boldsymbol{m}_0 \cdot \mathbf{N} \cdot \boldsymbol{m}_0 = 2\pi M_\xi^2(h - N_\zeta + N_\xi) + 2\pi M_\eta^2(h - N_\zeta + N_\eta) \tag{22}$$

where we have resolved $\boldsymbol{m}_0$ into the two components $\boldsymbol{M}_\xi$ and $\boldsymbol{M}_\eta$, along the ξ and η axes as previously defined.

The energy increase associated with a homogeneous rotation comes about by the excitation of $\boldsymbol{k}=0$ magnons, whence

$$E-E_0=n_0\hbar\omega(0) \tag{23}$$

and we therefore find

$$n_0=\frac{2\pi M_\xi{}^2(h-N_\zeta+N_\xi)+2\pi M_\eta{}^2(h-N_\zeta+N_\eta)}{\hbar\omega_M(h-N_\zeta+N_\xi)^{\frac{1}{2}}(h-N_\zeta+N_\eta)^{\frac{1}{2}}}. \tag{24}$$

An interesting special case arises when $N_\xi=N_\eta$, and

$$n_0=2\pi(M_\xi{}^2+M_\eta{}^2)/\hbar\omega_M\simeq(M_\zeta-M)/\gamma_0\hbar. \tag{25}$$

We thus find that equation (19) is valid for an ellipsoid of revolution as well as for an infinite medium.

The essential results of equations (16) and (24) are that *excitation of $\boldsymbol{k}=0$ magnons rotates the macroscopic magnetization, maintaining its magnitude, whereas excitation of $\boldsymbol{k}\neq 0$ magnons decreases the magnitude of the magnetization.*

3. RELATION OF DYNAMICAL EQUATION TO MAGNON TRANSITIONS

Consider, for definiteness, a system in equilibrium in the presence of an applied field. The number of $\boldsymbol{k}\neq 0$ magnons present at reasonably low temperatures is small and determines the thermal equilibrium value of the magnetization. The number of $\boldsymbol{k}=0$ magnons present is also very small, corresponding to the small thermal deviation of the direction of $\boldsymbol{M}$ from that of $\boldsymbol{H}$.

We suppose that the direction of the applied field is suddenly altered. Immediately thereafter the system finds itself with the same number of $\boldsymbol{k}\neq 0$ magnons, corresponding to the fact that the magnitude of $\boldsymbol{M}$ has not been changed. However, the number of $\boldsymbol{k}=0$ magnons has been suddenly increased to a very large number. If the new direction of the applied field defines a ζ-axis, then the number of $\boldsymbol{k}=0$ magnons follows from equation (24) or (25). As $M/\gamma_0\hbar$ is of the order of 10^{22}, the number of $\boldsymbol{k}=0$ magnons may be of the order of 10^{20} for small (1°) inclinations of the magnetization.

The large number of $\boldsymbol{k}=0$ magnons appears alarming when it is recalled that the spin-wave theory is a linearized theory valid only for small excitations. Fortunately, however, the number of $\boldsymbol{k}=0$ magnons may be uniquely large without invalidating the theory. That this is so may be seen as follows. The restriction to small numbers of magnons is necessary in spin-wave theory to prevent the formation of localized wave-packets of spin inversion, which would invert more than a single spin at a single lattice site. But the $\boldsymbol{k}=0$ magnons do not contribute to such localization because of their long wavelength. Therefore, large numbers of $k=0$ magnons do not invalidate the theory in the same critical fashion as do large numbers of other magnons.

However, at sufficiently large numbers of $\boldsymbol{k}=0$ magnons, a non-linear coupling with certain other magnons does become important. This non-linear coupling has been discussed by SUHL[12] and used to account for certain observed saturation effects at high power levels in ferromagnetic resonance. We shall assume that the number of $\boldsymbol{k}=0$ magnons, though large, is not sufficient to make the non-linear behavior important. If applied to ferromagnetic resonance, our dynamical equation will apply only to low or moderate power levels, at which saturation effects do not occur.

Returning to the system in which an applied field has suddenly been changed in direction, the system finds itself with a very large number of $\boldsymbol{k}=0$ magnons. If the spin-waves were absolutely true normal modes of the system, the number of magnons of each wave-vector would remain constant thereafter, and the macroscopic magnetization would precess with the frequency $\omega(0)$ of the $\boldsymbol{k}=0$ spin-wave. In actuality the system eventually comes into equilibrium with the applied field. The final number of $\boldsymbol{k}=0$ magnons returns to the initial equilibrium number, and if the number of $\boldsymbol{k}\neq 0$ magnons temporarily departs from its equilibrium number, it too returns. The energy represented in the large number of $\boldsymbol{k}=0$ magnons ultimately appears in the creation of phonons.

The dissipative process by which the system comes into equilibrium with an applied field consists of transitions among the magnons and phonons and, most particularly, of the destruction of $\boldsymbol{k}=0$ magnons.

We shall reserve for the next section a discussion of the dependence of the magnon transition

VOL. 4 1958

probabilities on the number of magnons, on the applied field, and on various intrinsic properties of the material. However, we can now see how the magnon transition probabilities are related to the parameters in the ferromagnetic dynamical equation.

We write the dynamical equation (1) in the more explicit form

$$\dot{\boldsymbol{M}}=\alpha\boldsymbol{M}-\gamma\boldsymbol{M}\times[\boldsymbol{H}_0-4\pi\boldsymbol{M}\cdot\mathbf{N}]-\lambda\boldsymbol{M}\times \times\{\boldsymbol{M}\times[\boldsymbol{H}_0-4\pi\boldsymbol{M}\cdot\mathbf{N}]\}. \tag{26}$$

Taking the scalar product with $\boldsymbol{M}/M$, we find

$$\dot{M}=\alpha M. \tag{27}$$

Differentiating equation (16) and equating to (27) then identifies α as

$$\alpha=-\frac{\gamma_0\hbar}{M}\dot{n}'. \tag{28}$$

To obtain a similar expression for λ, we write out equation (26) in components. To first order in M_ξ and M_η we find

$$\dot{M}_\xi=\alpha M_\xi-\frac{\gamma}{\gamma_0}\omega_M M_\eta[h-N_\zeta+N_\eta]+ +4\pi M^2\lambda M_\xi[h-N_\zeta+N_\xi] \tag{29}$$

$$\dot{M}_\eta=\alpha M_\eta+\frac{\gamma}{\gamma_0}\omega_M M_\xi[h-N_\zeta+N_\xi]+ +4\pi M^2\lambda M_\eta[h-N_\zeta+N_\eta] \tag{30}$$

and

$$\dot{M}_\zeta=\alpha M_\zeta. \tag{31}$$

But, from equation (24)

$$\dot{n}_0=\frac{4\pi M_\xi\dot{M}_\xi[h-N_\zeta+N_\xi]+4\pi M_\eta\dot{M}_\eta[h-N_\zeta+N_\eta]}{\hbar\omega_M[h-N_\zeta+N_\xi]^{\frac{1}{2}}[h-N_\zeta+N_\eta]^{\frac{1}{2}}}. \tag{32}$$

Inserting equations (29) and (30) into (32), we find

$$\frac{\dot{n}_0}{n_0}+\frac{2\gamma_0\hbar}{M}\dot{n}'=8\pi M^2\lambda\frac{M_\xi^2[h-N_\zeta+N_\xi]^2+M_\eta^2[h-N_\zeta+N_\eta]^2}{M_\xi^2[h-N_\zeta+N_\xi]+M_\eta^2[h-N_\zeta+N_\eta]}. \tag{33}$$

The loss parameter λ therefore does not depend only on n_0, n', $\dot{n}_0$, and $\dot{n}'$, but also explicitly on M_ξ and M_η; or, alternatively stated, λ varies during each precessional cycle. This point is a rather important one, and bears further consideration.

If the parameters α and λ were to be zero, the magnetization would precess in a cone. This cone is such that during a precession the energy, as given in equation (22), remains constant. The parameter λ describes a sort of viscous drag on the magnetization as it precesses around the cone. The value of λ is given by equation (33). The two terms of the left, as we shall subsequently see, are constant during a precessional cycle. But it is easily verified that the fraction on the right varies sinusoidally between $[h-N_\zeta+N_\xi]$ and $[h-N_\zeta+N_\eta]$, with a frequency which is twice the precessional frequency. The effect of this term is, then, to impart a sort of a "wobble" to the viscous drag or to modulate the viscous drag with twice the precessional frequency. If λ is small, the magnetization precesses many cycles in the time that the viscous drag exerts an appreciable effect, and the modulation is not important. Then the fraction on the right-hand side of equation (33) can be replaced by its average value. But if λ is large, so that the magnetization departs appreciably from its cone in less than a cycle, the modulating fraction in equation (33) becomes important and must be considered explicitly. We shall assume that the material has low loss, and we replace the fraction by its value averaged over a precessional cycle.

$$\frac{M_\xi^2[h-N_\zeta+N_\xi]^2+M_\eta^2[h-N_\zeta+N_\eta]^2}{M_\xi^2[h-N_\zeta+N_\xi]+M_\eta^2[h-N_\zeta+N_\eta]} \to\frac{[h-N_\zeta+N_\xi]+[h-N_\zeta+N_\eta]}{2} =h-\tfrac{3}{2}N_\zeta+\tfrac{1}{2}. \tag{34}$$

It should be noted that equation (34) is identically true if $N_\xi=N_\eta$, independently of the size of λ. Thus, this equation is approximately valid

either for small axial assymetry ($N_\xi \simeq N_\eta$) or for small values of λ.

We thus obtain

$$\lambda = \left(\frac{\dot{n}_0}{n_0} + \frac{2\gamma_0 \hbar}{M}\dot{n}'\right) \Big/ 8\pi M^2(h - \tfrac{3}{2}N_\zeta + \tfrac{1}{2}). \quad (35)$$

Equation (28) for α and the above equation for λ provide a connexion between the macroscopic dynamical equation and the underlying spin-wave representation.

The remaining macroscopic parameter γ essentially describes the angular velocity of the homogeneous precession, or the natural frequency of the $\boldsymbol{k} = 0$ magnons. The matrix elements which couple the $\boldsymbol{k} = 0$ magnons to other modes not only produce the transitions considered above, but they also slightly shift the $\boldsymbol{k} = 0$ frequency. Thus we conclude that γ is a constant, γ_c, which is very close to γ_0, and which becomes equal to γ_0 for those materials for which α and λ vanish.

The ferromagnetic dynamical equation becomes

$$\dot{\boldsymbol{M}} = -\gamma_0 \hbar \dot{n}' \frac{\boldsymbol{M}}{M} - \gamma_c \boldsymbol{M} \times \boldsymbol{H} - \frac{\dot{n}_0/n_0 + 2\gamma_0 \hbar \dot{n}'/M}{8\pi M^2(h - \tfrac{3}{2}N_\zeta + \tfrac{1}{2})} \boldsymbol{M} \times (\boldsymbol{M} \times \boldsymbol{H}). \quad (36)$$

4. THE DISSIPATIVE PROCESSES

To proceed further with the dynamical equation (36), we must appeal to some model to determine the functional dependence of $\dot{n}_0$ and $\dot{n}'$ on the state of the system, on the applied field, and on the intrinsic properties of the material.

Three principal types of transitions may play roles in the dissipative process. (1) The first type of magnon transition consists in the destruction of a $\boldsymbol{k} = 0$ magnon without the creation of another magnon. The $\boldsymbol{k} = 0$ magnon may create a phonon directly, scatter a phonon, interact with a conduction electron, or partake in some multiple magnon-phonon scattering process. (2) The second type of magnon transition consists in the destruction of a $\boldsymbol{k} = 0$ magnon with the associated creation of another magnon. Again, the process may involve phonons, conduction electrons, other magnons, etc. (3) The third type of magnon transition consists in the destruction of magnons with $\boldsymbol{k} \neq 0$.

Early theories of the magnon transitions concentrated on the direct magnon–phonon interaction. These theories[1] had great difficulty in accounting for ferromagnetic dissipation because of the necessity of simultaneously conserving energy– and crystal-momentum, despite the fact that the energy-momentum relations for magnons and phonons are different.

More recently the importance of processes which conserve energy but not momentum has been proposed by CLOGSTON *et al.*[2] These authors point out that the lack of strict periodicity in a disordered alloy or ferrite destroys the requirement of conservation of crystal-momentum, and they have also demonstrated that the alteration of the spin-wave spectrum by the demagnetizing fields makes the $\boldsymbol{k} = 0$ spin-wave degenerate with other spin-waves, so that energy can be conserved in magnon–magnon transitions.

The theories of the various types of magnon transitions have certain characteristics in common, which we shall exhibit by reviewing the description of several representative processes.

The momentum-non-conserving process in which a $\boldsymbol{k} = 0$ magnon is destroyed and a phonon is created is described by an interaction term in the Hamiltonian, of the form

$$\mathscr{H}_{0\boldsymbol{\sigma}} = A(\boldsymbol{\sigma})[b_{\boldsymbol{\sigma}}^{+} S_0^{-} + b_{\boldsymbol{\sigma}}^{-} S_0^{+}] \quad (37)$$

where S_0^- and S_0^+ are destruction and creation operators for $\boldsymbol{k} = 0$ magnons, and $b_{\boldsymbol{\sigma}}^+$ and $b_{\boldsymbol{\sigma}}^-$ are creation and destruction operators for phonons of wave-vector $\boldsymbol{\sigma}$. The coefficient $A(\boldsymbol{\sigma})$ is a measure of the coupling between the $\boldsymbol{k} = 0$ magnons and the phonons; it depends on the disorder in the lattice. The first term in equation (37) represents the destruction of a magnon and the creation of a phonon, and the second term represents the creation of a magnon and the destruction of a phonon. The symmetric combination is dictated by the Hermitian nature of the Hamiltonian.

The indirect momentum-non-conserving process destroys a $\boldsymbol{k} = 0$ magnon and creates a $\boldsymbol{k} \neq 0$ magnon; then, as a second step, the $\boldsymbol{k} \neq 0$ magnons are destroyed with the creation of phonons. The relevant interaction terms in the Hamiltonian are, for the first transition,

$$\mathscr{H}_{0\boldsymbol{k}} = B(\boldsymbol{k})[S_{\boldsymbol{k}}^{+} S_0^{-} + S_{\boldsymbol{k}}^{-} S_0^{+}] \quad (38)$$

VOL. 4 1958

and for the second transition

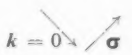
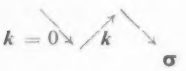
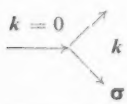
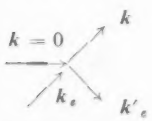
$$\mathscr{H}_{k\sigma} = C(\boldsymbol{k}, \boldsymbol{\sigma})[b_{\sigma}{}^{+}S_{\boldsymbol{k}}{}^{-} + b_{\sigma}{}^{-}S_{\boldsymbol{k}}{}^{+}]. \qquad (39)$$

Here, again, $B(\boldsymbol{k})$ and $C(\boldsymbol{k}, \boldsymbol{\sigma})$ are functions depending upon the magnitude of the relevant interactions.

The relations given above for the direct and indirect transitions are summarized in the first and second columns of Table 1. In that table we also give two representative examples of momentum-conserving dissipative processes. The first of these[1] consists in the scattering of a magnon from the state ($\boldsymbol{k} = 0$) to some other state $\boldsymbol{k}$, with the emission of a phonon of wave-vector $\boldsymbol{\sigma}$. The second representative momentum-conserving process cited in Table 1 is that arising from the interaction of the ferromagnetic $3d$ electrons with the conduction $4s$ electrons. This effect has been treated most recently by MITCHELL.[13] A $\boldsymbol{k} = 0$ magnon is scattered to a state $\boldsymbol{k}$, while a conduction electron with wave-vector $\boldsymbol{k}_e$ is scattered to a state $\boldsymbol{k}_e - \boldsymbol{k}$. The relevant interaction Hamiltonian for this process, as given in Table 1, is a product of magnon destruction and creation operators $S_{\boldsymbol{k}}{}^{\pm}S_0{}^{\mp}$ and a factor ϕ_e referring to the conduction electrons, given explicitly by MITCHELL.

The momentum-conserving processes are higher-order processes and have relatively small probability in disordered samples. They are included in Table 1 to stress that the conclusions we

Table 1. *Summary of ferromagnetic dissipative processes*

Transition	Hamiltonian	Net transition probability
I—*Momentum non-conserving (disordered samples)*		
Direct $(\boldsymbol{k} = 0) \rightarrow$ lattice  	$\mathscr{H}_{0\sigma} \simeq b_{\sigma}{}^{\pm}S_0{}^{\mp}$	$P_{0\sigma} \simeq n_0 - n_{\sigma} \doteq n_0$
Indirect $(\boldsymbol{k} = 0) \rightarrow (\boldsymbol{k} \neq 0) \rightarrow$ lattice  	$\mathscr{H}_{0k} \simeq S_{\boldsymbol{k}}{}^{\pm}S_0{}^{\mp}$ $\mathscr{H}_{k\sigma} \simeq b_{\sigma}{}^{\pm}S_{\boldsymbol{k}}{}^{\mp}$	$P_{0k} \simeq n_0 - n_{\boldsymbol{k}} \doteq n_0$ $P_{k\sigma} \simeq n_{\boldsymbol{k}} - n_{\sigma}$
II—*Momentum conserving (ordered samples)*		
Magnon–phonon scattering $(\boldsymbol{k} = 0) \rightarrow (\boldsymbol{k} \neq 0) + \boldsymbol{\sigma}$  	$\mathscr{H}_{0k\sigma} \simeq b_{\sigma}{}^{\pm}S_{\boldsymbol{k}}{}^{\pm}S_0{}^{\mp}$ $(\boldsymbol{\sigma} = -\boldsymbol{k})$	$P_{0k\sigma} \simeq (n_{\sigma} + n_{\boldsymbol{k}} + 1)n_0 - n_{\sigma}n_{\boldsymbol{k}} \doteq n_0$
Magnon-conduction electron scattering  	$\mathscr{H}_{0k,e} \simeq \phi_e S_{\boldsymbol{k}}{}^{\pm}S_0{}^{\mp}$	$P_{0k,e} \doteq n_0$

shall draw have a wide generality and are not predicated upon the assumption of a particular type of dissipative process.

The transition probability for each process is proportional to the square of the matrix element of the interaction Hamiltonian, and the squares of matrix elements of creation and destruction operators can be written in terms of the numbers of magnons or phonons.

As a specific example we shall consider the direct momentum-non-conserving process. The interaction Hamiltonian $b_\sigma^{\pm}S_0^{\mp}$ has a matrix element corresponding to the destruction of a $\boldsymbol{k}=0$ magnon and the creation of a phonon

$$|\langle b_\sigma^{+}S_0^{-}\rangle|^2 \simeq (n_\sigma+1)n_0. \tag{40}$$

It also has a matrix element corresponding to the creation of a $\boldsymbol{k}=0$ magnon and the destruction of a phonon

$$|\langle b_\sigma^{-}S_0^{+}\rangle|^2 \simeq n_\sigma(n_0+1). \tag{41}$$

The net probability of destruction of a $k=0$ magnon is the difference of the two probabilities in equations (40) and (41)

$$P_{0\sigma} \simeq (n_\sigma+1)n_0-n_\sigma(n_0+1)=n_0-n_\sigma. \tag{42}$$

Finally, we recall that the number n_σ of phonons with wave-vector $\boldsymbol{\sigma}$ is a small equilibrium number (of the order of one), whereas n_0 is an enormously large number. Hence we can ignore n_σ in comparison with n_0 in equation (42) and *the probability of destruction of a* $\boldsymbol{k}=0$ *magnon becomes proportional to the number of* $\boldsymbol{k}=0$ *magnons.*

$$P_{0\sigma} \simeq n_0. \tag{43}$$

The indirect non-momentum-conserving process also merits separate and explicit consideration. As indicated in the third column of Table 1, the net probability of destruction of a $\boldsymbol{k}=0$ magnon (and creation of a $\boldsymbol{k}\neq 0$ magnon) is given by equations analogous to (42) and (43)

$$P_{0k} \simeq (n_k+1)n_0-n_k(n_0+1)=n_0-n_k \simeq n_0. \tag{44}$$

The second step of the process, in which the $\boldsymbol{k}\neq 0$ magnon is destroyed and a phonon created, has a probability given by

$$P_{k\sigma} \simeq (n_\sigma+1)n_k-n_\sigma(n_k+1)=n_k-n_\sigma. \tag{45}$$

Both n_k and n_σ are small numbers and neither can be neglected relative to the other.

In both equations (43) and (44), and in every other reasonable model, the probability of destruction of a $\boldsymbol{k}=0$ magnon is proportional to n_0. The process (43) is of the first type alluded to at the beginning of this section; it does not produce a magnon of $\boldsymbol{k}\neq 0$. But the process of equation (44) is of the second type; it is associated with the production of $\boldsymbol{k}\neq 0$ magnons. We thus find that, in general

$$\dot{n}_0 = -(\lambda_{0\sigma}+\lambda_{0k})n_0, \tag{46}$$

where $\lambda_{0\sigma}$ refers to the first type of process and λ_{0k} refers to the second type of process.

The nature of the parameters $\lambda_{0\sigma}$ and λ_{0k} can be understood in terms of the standard perturbation-theory formulae. Characteristically $\lambda_{0\sigma}$ is the product of two factors. The first factor is the square of a matrix element connecting the singly-occupied $\boldsymbol{k}=0$ spin-wave state with some final phonon (or "flipped" conduction electron) states. The second factor is a density-in-energy of the final phonon or conduction electron states.

Similarly λ_{0k} is characteristically the product of two factors. Again, the first factor is the square of a matrix element connecting a $\boldsymbol{k}=0$ and a $\boldsymbol{k}\neq 0$ magnon, and the second factor is a density-in-energy of those $\boldsymbol{k}\neq 0$ magnons which are degenerate with $\boldsymbol{k}=0$.

The rate of change of n' arises from two effects. The destruction of $\boldsymbol{k}=0$ magnons with the creation of $\boldsymbol{k}\neq 0$ magnons contributes a term $\lambda_{0k}n_0$, and the destruction of $\boldsymbol{k}\neq 0$ magnons by effects such as equation (45) contributes a negative term.

In equation (45) the number of phonons n_σ is the equilibrium number at the temperature in question. This number is equal to the equilibrium number n_T' of $\boldsymbol{k}\neq 0$ magnons, with which the phonons are degenerate. We now assume that all such processes contribute a term to $\dot{n}'$ of the form $-\lambda_{k\sigma}(n'-n_T')$. In adopting this assumption, we recognize that $\lambda_{k\sigma}$ is by no means as fundamental a parameter as $\lambda_{0\sigma}$ or λ_{0k}. Both $\lambda_{0\sigma}$ and λ_{0k} are concerned with transitions of a single type of magnon: the $\boldsymbol{k}=0$ magnon. But $\lambda_{k\sigma}$ represents a complicated average over transition probabilities of magnons with many different $\boldsymbol{k}$ values. Nevertheless the assumption of a net transition probability proportional to $(n'-n'_T)$ is probably justified,

VOL. 4 1958

partially on the basis of equation (45) and partially by the realization that n' is always very close to n'_T. This assumption, of course, is just the "spin-lattice relaxation time" assumption of the BLOCH–BLOEMBERGEN theory.[5]

We then have

$$\dot{n}' = \lambda_{0k} n_0 - \lambda_{k\sigma}(n' - n'_T) \tag{47}$$

or by equation (16)

$$\gamma_0 \hbar \dot{n}' = \lambda_{0k} \gamma_0 \hbar n_0 + \lambda_{k\sigma}(M - M_T), \tag{48}$$

where M_T is the equilibrium value of M at the temperature T.

Inserting equations (46) and (48) into (36), we find the basic form of the dynamical equation

$$\dot{\boldsymbol{M}} = -[\lambda_{0k}\gamma_0\hbar n_0 + \lambda_{k\sigma}(M - M_T)]\frac{\boldsymbol{M}}{M} - \gamma_e \boldsymbol{M} \times \boldsymbol{H} + \\ + \left[\frac{\lambda_{0\sigma} M + \lambda_{0k}(M - 2\gamma_0\hbar n_0) - 2\lambda_{k\sigma}(M - M_T)}{8\pi M^3 (h - \frac{3}{2}N_\zeta + \frac{1}{2})}\right] \times \\ \times \boldsymbol{M} \times (\boldsymbol{M} \times \boldsymbol{H}). \tag{49}$$

In this equation the quantity n_0 depends upon the direction of the magnetization, and is to be computed by equation (24) or (25); the effective demagnetization coefficient N_ζ depends upon the direction and magnitude of the applied field, in the manner defined by equations (7) and (13); and the three quantities λ_{0k}, $\lambda_{k\sigma}$, and $\lambda_{0\sigma}$ depend in a yet-unspecified way on the field, shape, size, order, etc.

It is clear that equation (49) is not yet in an explicit and usable form. The virtue of the equation is that the parameters λ_{0k}, $\lambda_{k\sigma}$, and $\lambda_{0\sigma}$ have definite quantum-mechanical meaning. Equation (49) therefore establishes a definite connection between the macroscopic dynamical equation and the underlying quantum-mechanical processes, and it provides the framework for theoretical and experimental investigation.

Two special cases of equation (49) are of particular interest. First, suppose that the $\boldsymbol{k} = 0$ magnons are not coupled with other magnons, so that λ_{0k} vanishes. Then if $M = M_T$ initially, it thenceforth maintains this magnitude, and equation (49) becomes

$$\dot{\boldsymbol{M}} = -\gamma_e \boldsymbol{M} \times \boldsymbol{H} + \left[\frac{\lambda_{0\sigma}}{8\pi M^2 (h - \frac{3}{2}N_\zeta + \frac{1}{2})}\right] \times \\ \times \boldsymbol{M} \times (\boldsymbol{M} \times \boldsymbol{H}). \tag{50}$$

A second special case of interest arises when the coupling of the $\boldsymbol{k} \neq 0$ magnons to the phonons is very much stronger than the coupling of the $\boldsymbol{k} = 0$ magnons to the $\boldsymbol{k} \neq 0$ magnons. That is

$$\lambda_{k\sigma} \gg \lambda_{0k}. \tag{51}$$

Then, by equation (47) or (48), we see that the phonons "thermostat" the number of $\boldsymbol{k} \neq 0$ magnons, so that the instantaneous number of $\boldsymbol{k} \neq 0$ magnons is determined by

$$M - M_T \doteq -\frac{\lambda_{0k}}{\lambda_{k\sigma}} \gamma_0 \hbar n_0 \doteq 0. \tag{52}$$

That is, for sufficiently large $\lambda_{k\sigma}$ the phonon field ensures that the magnetization maintains very nearly its equilibrium magnitude. Equation (49) then becomes

$$\dot{\boldsymbol{M}} = -\gamma_e \boldsymbol{M} \times \boldsymbol{H} + \left[\frac{\lambda_{0\sigma} + \lambda_{0k}}{8\pi M^2 (h - \frac{3}{2}N_\zeta + \frac{1}{2})}\right] \times \\ \boldsymbol{M} \times (\boldsymbol{M} \times \boldsymbol{H}). \tag{53}$$

Equations (50) and (53) both describe purely rotational response, and they have a comparatively simple form. However, even in these cases, the functional form of $\lambda_{0\sigma}$ and λ_{0k} must be investigated. As an illustration of the manner in which a detailed model may yield more specific information we shall consider a particular mechanism in Section 5.

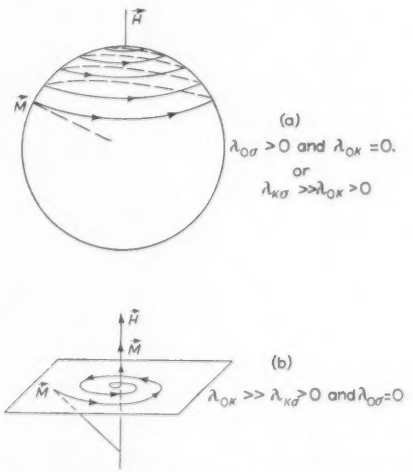


Fig. 4. Representative paths of magnetization.

The paths of the magnetization for several specific dissipative mechanisms are illustrated in Fig. 4. The response is shown for an infinite sample, in which the direction of the internal field remains constant; in a finite ellipsoidal sample, the demagnetization causes the internal field to change direction continually and obscures the pictorial representation of the response. In Fig. 4(a) the response is shown for the case in which $\lambda_{0\sigma} > 0$ and $\lambda_{0k} = 0$. The number of $\boldsymbol{k} \neq 0$ magnons, and consequently the magnitude of the magnetization, is conserved, and the magnetization spirals into the field direction on the surface of a sphere. The path shown in Fig. 4(a) also represents the response if $\lambda_{k\sigma} \gg \lambda_{0k} > 0$. In that case the lattice "thermostats" the number of $\boldsymbol{k} \neq 0$ magnons, and thereby maintains the magnitude of the magnetization at its equilibrium value.

In Fig. 4(b) the magnetization path is shown for the case $\lambda_{0k} \gg \lambda_{k\sigma} > 0$ and $\lambda_{0\sigma} = 0$. The response then takes place in two stages. In the first stage the rapid $(\boldsymbol{k} = 0) \rightarrow (\boldsymbol{k} \neq 0)$ transitions occur. Such transitions do not change the total number of magnons and, consequently, do not change the component of $\boldsymbol{M}$ along $\boldsymbol{H}$; the magnetization spirals into the field direction on the surface of a plane perpendicular to $\boldsymbol{H}$. In the second stage of the response, the slow $(\boldsymbol{k} \neq 0) \rightarrow (\boldsymbol{\sigma})$ transitions occur, and the magnetization increases in magnitude to its equilibrium value, "growing" along the field direction.

5. A SPECIAL CASE—THE CLOGSTON MECHANISM

In order to specialize the general dynamical equation in one illustrative case, we shall assume that the loss occurs purely by the CLOGSTON mechanism. That is, we assume that the destruction of $\boldsymbol{k} = 0$ magnons is always associated with the production of $\boldsymbol{k} \neq 0$ magnons, so that $\lambda_{0\sigma} = 0$. We further assume that $\lambda_{k\sigma} \gg \lambda_{0k}$, so that equation (53) holds. With both of these assumptions we find

$$\dot{\boldsymbol{M}} = -\gamma_e \boldsymbol{M} \times \boldsymbol{H} + \left[\frac{\lambda_{0k}}{8\pi M^2 (h - \frac{3}{2}N_\zeta + \frac{1}{2})}\right] \times \times \boldsymbol{M} \times (\boldsymbol{M} \times \boldsymbol{H}). \tag{54}$$

We now turn to the problem of calculating λ_{0k}. This calculation, of course, is independent of the assumptions leading to equation (54).

According to the definition in equation (46), λ_{0k} is the probability, per $\boldsymbol{k} = 0$ magnon, of the destruction of a $\boldsymbol{k} = 0$ magnon with the associated production of a $\boldsymbol{k} \neq 0$ magnon. It is the product of a matrix element squared and a density-in-energy of final stages.

Consider first the square of the matrix element connecting a $\boldsymbol{k} = 0$ magnon with a $\boldsymbol{k} \neq 0$ magnon. This matrix element vanishes unless $\omega(\boldsymbol{k}) = \omega(0)$. It also vanishes for every $\boldsymbol{k}$ in a perfectly periodic crystal. In a disordered crystal the magnitude of the matrix element depends upon the $\boldsymbol{k}^{\text{th}}$ Fourier amplitude of the disorder potential. In order to proceed in the simplest possible way, we now adopt the assumption that the lattice disorder has a completely "white" spectrum. That is, we assume that every Fourier amplitude of the disorder has equal magnitude. The matrix element is then independent of the final magnon wave-vector, except for the requirement of energy conservation.

The assumption of a white disorder spectrum is an analytic convenience made for the purpose of obtaining a preliminary theory. CLOGSTON *et al.*[2] introduce a pseudo-dipolar interaction and assume certain correlations between spins, so as to introduce implicitly a different disorder spectrum. Nevertheless, their result and ours are very similar. Mr. E. PITTELLI and I are presently investigating the role of short-range order in determining the disorder spectrum, so as to obtain a relationship between ferromagnetic dissipation and heat treatment.

Under the assumption of constancy of the matrix element, λ_{0k} simply becomes proportional to the density-in-energy of the final states. This density of states can be computed as follows.

The wave-vectors of the magnons degenerate with the $\boldsymbol{k} = 0$ magnons may be found by equating $\omega(\boldsymbol{k})$, as given in equation (9), to $\omega(0)$, as given in equation (10). We thus find

$$(h - N_\zeta + \kappa^2)(h - N_\zeta + \kappa^2 + \sin^2\theta) = (h - N_\zeta + N_\xi)(h - N_\zeta + N_\eta) = \omega^2(0)/\omega_M{}^2. \tag{55}$$

These degenerate spin-wave states are the final states, the density of which we wish to calculate. They are the states shown in Fig. 5 for two different shapes of sample. It is clear from that figure that the number of these states does depend on the

VOL. 4 1958

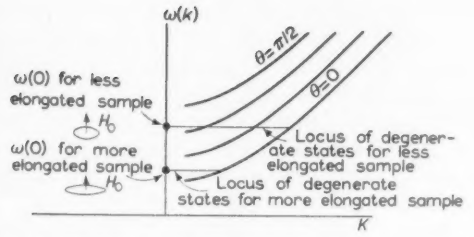


FIG. 5. Effect of sample shape on magnon degeneracy.

sample shape, illustrating the source of the shape-dependence of λ_{0k}.

Another representation of the states degenerate with the $\boldsymbol{k} = 0$ state is shown in Fig. 6. Equation

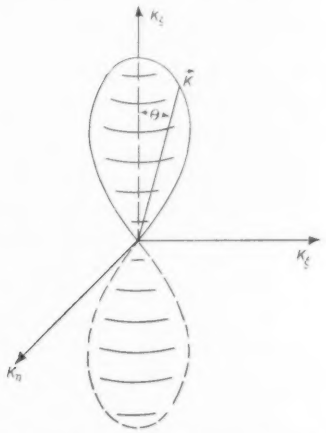


FIG. 6. The locus of magnons degenerate with the zero-wave-vector magnons.

(55) is a relation between k and θ, which determines a surface in reciprocal space. Such a surface is shown in Fig. 6. All wave-vectors terminating on this surface represent spin-wave states degenerate with the $\boldsymbol{k} = 0$ state.

The density-in-energy of the final states is proportional to the volume in reciprocal space between two surfaces corresponding to unit energy difference. That is, if two surfaces are drawn, one for a given value of $\omega(0)$ and one for $\omega(0)+d\omega(0)$, the density-in-energy is proportional to the volume between them, divided by $\hbar d\omega(0)$.

Thus, if $V(\omega(0))$ is the volume in reciprocal space contained within the surface shown in Fig. 6, then

$$\lambda_{0k} = \lambda_{0k}{}^{0}\frac{dV(\omega(0))}{\hbar d\omega(0)} \tag{56}$$

where $\lambda_{0k}{}^{0}$ is a proportionality constant.

The integral involved in equation (56) has been reduced to tabulated functions by PITTELLI.[14] Let κ (0) be the value of κ for $\theta = 0$, then

$$V = 4/3\pi\kappa^3(0) - 4\pi\int_0^{\kappa(0)}\int_{\theta(\kappa)}^{\pi/2}\kappa^2\sin\theta\, d\theta d\kappa. \tag{57}$$

Carrying out the integral over θ and introducing the new variable ϵ by the definition

$$\epsilon = h - N_\zeta + \kappa^2 \tag{58}$$

we find

$$V = 4/3\pi\kappa^3(0) - 2\pi\times$$
$$\times\int_{h-N_\zeta}^{\omega(0)/\omega_M}(\epsilon - h + N_\zeta)^{\frac{1}{2}}\left[1+\epsilon-\frac{\omega^2(0)}{\omega_M{}^2\epsilon}\right]^{\frac{1}{2}}d\epsilon. \tag{59}$$

Then, from equation (56)

$$\lambda_{0k} = \lambda_{0k}{}^{0}\frac{2\pi\omega(0)}{\omega^2{}_M}\int_{h-N_\zeta}^{\omega(0)/\omega_M}\frac{(\epsilon-h+N_\zeta)^2}{\epsilon^{\frac{1}{2}}(\epsilon-b)^{\frac{1}{2}}(\epsilon-d)^{\frac{1}{2}}}d\epsilon \tag{60}$$

where

$$b = -\tfrac{1}{2}+\tfrac{1}{2}(1+4\omega^2(0)/\omega_M{}^2)^{\frac{1}{2}} \tag{61}$$

and

$$d = -\tfrac{1}{2}-\tfrac{1}{2}(1+4\omega^2(0)/\omega_M{}^2)^{\frac{1}{2}}. \tag{62}$$

The magnitudes of these quantities are such that

$$\frac{\omega(0)}{\omega_M} > h - N_\zeta > b > 0 > d. \tag{63}$$

The integral in equation (60) is an elliptic integral which can be reduced to standard form.[15] The evaluated integral is

$$\lambda_{0k} = \lambda_{0k}{}^{0}\frac{\pi\omega(0)}{\omega_M{}^2}[h-N_\zeta-b]\times$$
$$\times[h^2(1+4\omega^2(0)/\omega_M{}^2)]^{-\frac{1}{2}}[\Pi(\phi,\beta^2,l)-F(\phi,l)] \tag{64}$$

where $\Pi(\phi, \beta^2, l)$ is Legendre's elliptic integral of

the third kind, $F(\phi, l)$ is Legendre's elliptic integral of the second kind, and

$$\sin\phi = \left[\frac{(\omega(0)/\omega_M - h + N_\zeta)(1+4\omega(0)/\omega_M)^{\frac{1}{2}}}{(h-N_\zeta-d)(\omega(0)/\omega_M - b)}\right]^{\frac{1}{2}} \tag{65}$$

$$\beta^2 = (h-N_\zeta-d)(1+4\omega^2(0)/\omega_M{}^2)^{-\frac{1}{2}} \tag{66}$$

$$l^2 = \frac{b(h-N_\zeta-d)}{h-N_\zeta}(1+4\omega^2(0)/\omega_M{}^2)^{-\frac{1}{2}}. \tag{67}$$

If desired, the elliptic integral of the third kind can be further reduced

$$\Pi(\phi,\beta^2,l) = \frac{-\beta}{(\beta^2-1)^{\frac{1}{2}}(\beta^2-l^2)^{\frac{1}{2}}}F(\phi,l)Z(\sin^{-1}1/\beta,l) -$$

$$-\tfrac{1}{2}\ln\frac{H[F(\sin^{-1}1/\beta,l)+F(\phi,l)]}{H[F(\sin^{-1}1/\beta,l)-F(\phi,l)]}, \tag{68}$$

where Z is the Jacobian zeta function and H is the Jacobian theta function.

By the use of these equations, λ_{0k} can be evaluated, using standard tables of functions. As an example of this procedure, PITTELLI has calculated the shape-dependence of λ_{0k} for an ellipsoid of revolution, assuming a value of $h = 1{\cdot}7$. The results of this calculation are shown in Fig. 7. The diminution of the loss for thin samples, predicted qualitatively in Fig. 5, is shown explicitly in Fig. 7.

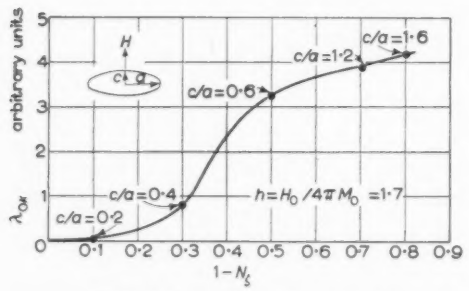


FIG. 7. Shape-dependence of λ_{0k}.

The field-dependence of the dynamical equation resides in equation (64), as does the shape-dependence. If equation (64), for λ_{0k}, is inserted into the dynamical equation (54), we obtain an explicit dynamical equation for rotational response. It is, of course, uncertain as to whether the underlying assumptions on the relative magnitudes of $\lambda_{0\sigma}$, λ_{0k}, and $\lambda_{k\sigma}$ are valid in any particular disordered material. This question demands experimental investigation, and theoretical examination of $\lambda_{0\sigma}$ and $\lambda_{k\sigma}$ in the same manner as that in which λ_{0k} has been treated here.

Another interesting effect can be understood in terms of a size effect on λ_{0k}, rather than a shape effect. Dr. HENRY BELSON has observed[16] that the ferromagnetic line width in thin Permalloy films increases with thickness (by a factor of the order of five for thicknesses between 1000 and 5000 Å) when the applied field is *parallel* to the plane of the film. Now as the film thickness goes to zero $\omega(0)$ approaches the bottom of the spectrum in Fig. 5, if the film is perpendicularly magnetized. But if the film is magnetized parallel to the plane, $\omega(0)$ approaches the top of the spectrum, that is, it approaches the extrapolated intercept of the $\theta = \pi/2$ curve. Small changes in shape then cannot make large percentage changes in λ_{0k}, as is clear from Fig. 5, and therefore shape effects are unable to account for BELSON's observations. However, in Fig. 8 we show the spin-wave spectrum of

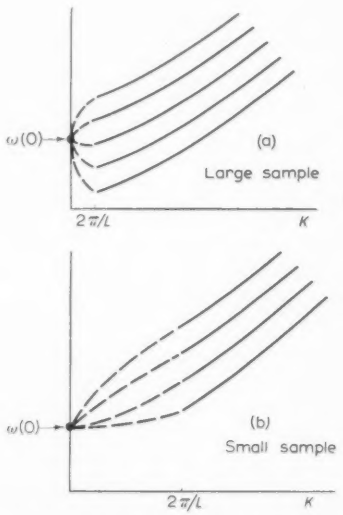


FIG. 8. Size-dependence of λ_{0k}. In the small sample the distortion of the parabolic spectrum occurs at such large k that few modes remain degenerate with the $k = 0$ mode.

VOL. 4 1958

two samples of the same shape but different sizes. The first spectrum is of a relatively large sample. The spin-waves follow parabolic curves to small $\boldsymbol{k}$, but when $\boldsymbol{k}$ becomes so small that $2\pi/k$ is comparable to the sample size, the spin-wave spectrum becomes distorted. This is the region of the WALKER[10] modes, indicated very schematically by the dotted curves in Fig. 8. In Fig. 8 we also show the spectrum for a very thin sample; here the distortion of the parabolic spectrum occurs at much larger $\boldsymbol{k}$, and may remove the degeneracy of the magnons responsible for the loss. Preliminary rough calculations of this effect, which PITTELLI and I have carried out, seem to give reasonable magnitudes for the size-dependence of the loss, but a more definite conclusion awaits further calculations.

APPENDIX

Spin-Wave Spectrum for Arbitrary Direction of the Applied Field

Consider an ellipsoidal sample with demagnetization $4\pi\mathbf{N}$. Let an external field $\boldsymbol{H}_0$ be applied in an arbitrary direction.

The magnetostatic energy is

$$E = -\boldsymbol{M}\cdot\boldsymbol{H}_0 + 2\pi\boldsymbol{M}\cdot\mathbf{N}\cdot\boldsymbol{M}. \tag{A-1}$$

The equilibrium direction of $\boldsymbol{M}$ is determined by the condition that the energy be minimum.

$$\delta E = -\delta\boldsymbol{M}\cdot[\boldsymbol{H}_0 - 4\pi\mathbf{N}\cdot\boldsymbol{M}] = 0. \tag{A-2}$$

Thus, if $\boldsymbol{\zeta}$ is a unit vector in the direction of the equilibrium magnetization and if H_e is the equilibrium value of the *internal* field,

$$\boldsymbol{H}_0 - 4\pi M_0\mathbf{N}\cdot\boldsymbol{\zeta} = H_e\boldsymbol{\zeta}, \quad H_e \geqslant 0. \tag{A-3}$$

We now let $\boldsymbol{M}$ deviate slightly from $M_0\boldsymbol{\zeta}$, defining a small vector $\boldsymbol{m}$ by

$$\boldsymbol{M} = M_0\boldsymbol{\zeta} + \boldsymbol{m} \tag{A-4}$$

where

$$\boldsymbol{m}\cdot\boldsymbol{\zeta} = 0. \tag{A-5}$$

More explicitly, we let $\boldsymbol{m}_0$ be the spatial average of $\boldsymbol{m}$, and $\boldsymbol{m}_1$ be the local deviation from this spatial average, so that

$$\boldsymbol{M} = M_0\boldsymbol{\zeta} + \boldsymbol{m}_0 + \boldsymbol{m}_1 \tag{A-6}$$

and

$$\boldsymbol{m}_0\cdot\boldsymbol{\zeta} = \boldsymbol{m}_1\cdot\boldsymbol{\zeta} = 0. \tag{A-7}$$

The deviation $\boldsymbol{m}_0$ will characterize the spin-wave of zero wave-vector, whereas $\boldsymbol{m}_1$ will characterize the resultant of all other spin-waves.

The local effective field will be

$$\boldsymbol{H} = \boldsymbol{H}_0 - 4\pi M_0\boldsymbol{\zeta}\cdot\mathbf{N} - 4\pi\boldsymbol{m}_0\cdot\mathbf{N} + \\ + \boldsymbol{H}_d(\boldsymbol{m}_1) + \boldsymbol{H}_{ex}(\boldsymbol{m}_1). \tag{A-8}$$

The first term here is the applied field, the second is the demagnetizing field due to the component $M_0\boldsymbol{\zeta}$, and the third term is the demagnetizing field due to $\boldsymbol{m}_0$. The fourth term in equation (A-8) is the demagnetizing field due to the magnetization $\boldsymbol{m}_1$; because of the spatial variation of $\boldsymbol{m}_1$, this term must be computed with careful attention to the local boundary conditions.[12] The final term in equation (A-8) is the exchange field, again arising from the spatial variations of $\boldsymbol{m}_1$.

We insert the above equations into the basic dynamical equation

$$\dot{\boldsymbol{M}} = -\gamma_0\boldsymbol{M}\times\boldsymbol{H} \tag{A-9}$$

in which we ignore dissipative terms because we are seeking normal modes. Then

$$\dot{\boldsymbol{m}} = -\gamma_0(M_0\boldsymbol{\zeta} + \boldsymbol{m})\times[\boldsymbol{H}_0 - 4\pi M_0\boldsymbol{\zeta}\cdot\mathbf{N} - \\ - 4\pi\boldsymbol{m}_0\cdot\mathbf{N} + \boldsymbol{H}_d(\boldsymbol{m}_1) + \boldsymbol{H}_{ex}(\boldsymbol{m})]. \tag{A-10}$$

Expanding, and keeping only terms which are first-order in the small quantity $\boldsymbol{m}$, we obtain

$$\dot{\boldsymbol{m}} = -\gamma_0 M_0\boldsymbol{\zeta}\times \\ \times[-4\pi\boldsymbol{m}_0\cdot\mathbf{N} + \boldsymbol{H}_d(\boldsymbol{m}_1) + \boldsymbol{H}_{ex}(\boldsymbol{m})] - \\ - \gamma_0[\boldsymbol{H}_0\cdot\boldsymbol{\zeta} - 4\pi M_0\boldsymbol{\zeta}\cdot\mathbf{N}\cdot\boldsymbol{\zeta}](\boldsymbol{m}\times\boldsymbol{\zeta}). \tag{A-11}$$

Let $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ be two vectors which, with $\boldsymbol{\zeta}$, form an orthonormal set. Define a tensor $\mathbb{N}$ by

$$\mathbb{N} = \mathbf{N} - (\boldsymbol{\zeta}\cdot\mathbf{N}\cdot\boldsymbol{\xi})(\boldsymbol{\xi}\boldsymbol{\zeta} + \boldsymbol{\zeta}\boldsymbol{\xi}) - \\ - (\boldsymbol{\zeta}\cdot\mathbf{N}\cdot\boldsymbol{\eta})(\boldsymbol{\eta}\boldsymbol{\zeta} + \boldsymbol{\zeta}\boldsymbol{\eta}) \tag{A-12}$$

where $(\boldsymbol{\xi}\boldsymbol{\zeta})$ is the dyadic product of $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$, etc. As $\mathbb{N}$ is obviously symmetric, it possesses three orthogonal *eigen*-axes. One of these is the ζ-axis, and it is easy to corroborate that

$$\boldsymbol{\zeta}\cdot\mathbb{N} = (\boldsymbol{\zeta}\cdot\mathbf{N}\cdot\boldsymbol{\zeta})\boldsymbol{\zeta}. \tag{A-13}$$

Equation (A-11) is clearly unaltered if we replace $\mathbf{N}$ by $\mathbb{N}$, so that we can write

$$\dot{\boldsymbol{m}} = -\gamma_0 M_0\boldsymbol{\zeta}\times \\ \times[-4\pi\boldsymbol{m}_0\cdot\mathbb{N} + \boldsymbol{H}_d(\boldsymbol{m}_1) + \boldsymbol{H}_{ex}(\boldsymbol{m})_1] - \\ - \gamma_0[\boldsymbol{H}_0\cdot\boldsymbol{\zeta} - 4\pi M_0\boldsymbol{\zeta}\cdot\mathbb{N}\cdot\boldsymbol{\zeta}](\boldsymbol{m}\times\boldsymbol{\zeta}). \tag{A-14}$$

But this equation is precisely the equation which we would write for a hypothetical sample with the demagnetization tensor $\mathbb{N}$ and with the field $\boldsymbol{H}_0\cdot\boldsymbol{\zeta}$ *applied along a principal axis*. This case has been treated by

ANDERSON and SUHL,[8] and we can merely adopt their solution of equation (A-14). The result is given in equations (8)—(13) of this paper.

The associated tensor $\mathbb{N}$ can be written in a convenient form. Consider first the symmetric dyadic tensor $(H_0\zeta+\zeta H_0)$. We find

$$(H_0\zeta+\zeta H_0)=(H_0\cdot\xi)(\xi\zeta+\zeta\xi)+$$
$$+(H_0\cdot\eta)(\eta\zeta+\zeta\eta)+2(H_0\cdot\zeta)(\zeta\zeta). \quad \text{(A-15)}$$

But, by equation (A-3), it follows that

$$H_0\cdot\xi=4\pi M_0\xi\cdot\mathbf{N}\cdot\zeta \quad \text{(A-16)}$$

and

$$H_0\cdot\eta=4\pi M_0\eta\cdot\mathbf{N}\cdot\zeta. \quad \text{(A-17)}$$

Equations (A-12), (A-15), and (A-16) now permit us to write

$$\mathbb{N}=\mathbf{N}-(h_0\zeta+\zeta h_0)+2(h_0\cdot\zeta)(\zeta\zeta) \quad \text{(A-18)}$$

where

$$h_0=H_0/4\pi M_0. \quad \text{(A-19)}$$

We thus have shown that the spin-wave spectrum in a sample $\mathbf{N}$ with applied field H_0 is equal to that in a hypothetical sample $\mathbb{N}$ with applied field $H_0\cdot\zeta$ *applied along the principal axis* ζ.

REFERENCES

1. KITTEL C. and ABRAHAMS E. *Rev. Mod. Phys.* **25,** 233 (1953).
2. CLOGSTON A. M., SUHL H., WALKER L. R. and ANDERSON P. W. *J. Phys. Chem. Solids* **1,** 129 (1956).
3. LANDAU L. and LIFSHITZ E. *Phys. Z. Sowjet.* **8,** 153 (1935).
4. GILBERT T. A. Armour Research Institute unpublished reports.
5. BLOEMBERGEN N. *Phys. Rev.* **78,** 572 (1950).
   WANGSNESS R. K. and BLOCH F. *Phys. Rev.* **89,** 728 (1953).
   WANGSNESS R. K. *Phys. Rev.* **98,** 927 (1955).
   BLOCH F. *Phys. Rev.* **102,** 104 (1956); **105,** 1206 (1957).
6. BLOCH F. *Z. Phys.* **61,** 206 (1930); **74,** 295 (1932).
7. HERRING C. and KITTEL C. *Phys. Rev.* **81,** 869 (1951).
8. ANDERSON P. W. and SUHL H. *Phys. Rev.* **100,** 1788 (1955).
9. DYSON F. J. *Phys. Rev.* **102,** 1217 (1956).
10. WALKER L. R. *Phys. Rev.* **105,** 390 (1957).
11. KEFFER F., KAPLAN H. and YAFET Y. *Amer. J. Phys.* **21,** 250 (1953).
12. SUHL H. *J. Phys. Chem. Solids* **1,** 209 (1956).
13. MITCHELL A. H. *Phys. Rev.* **105,** 1439 (1957).
14. PITTELLI E. Private communication.
15. BYRD and FRIEDMAN *Handbook of Elliptic Integrals.* Berlin, Springer (1954).
16. BELSON H. Private communication.

VOL. 4 1958

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 271–277.

SUSCEPTIBILITY OF THE THREE-LEVEL MASER

A. M. CLOGSTON

Bell Telephone Laboratories, Inc., Murray Hill, N.J.

(*Received* 28 *August* 1957)

Abstract—The quantum-mechanical equations of motion of a three-level maser, including the off-diagonal terms, are solved approximately to obtain the susceptibility χ presented by the paramagnetic material at the frequency $(1/h)(E_3-E_2)$. The effect of the cavity reaction at the frequency $(1/h)(E_2-E_1)$ is considered. The line shape is shown to be drastically altered for large-amplitude driving fields.

1. INTRODUCTION

THE great interest that has arisen in recent months in the field of maser oscillation and amplification has resulted in a variety of theoretical treatments of the dynamics of maser operation.[1–4] One particularly systematic way of approaching the maser consists in describing the response to a radiation field of the solid or gas involved by means of a frequency-dependent susceptibility. The ammonia maser, a two-level system, has been discussed in this way by GORDON *et al.*[1] If the susceptibility χ is written in the usual way as $\chi'-i\chi''$, where χ' describes the in-phase and χ'' the out-of-phase response of the system, the maser principle can be expressed by the statement that χ' and χ'' have the opposite sign to that encountered in a passive system. The negative value of χ'' presented to the radiation field counteracts the losses present in the electromagnetic system and thereby leads to regenerative amplification or oscillation. In the case of the ammonia maser, the change of sign arises, of course, from the experimental arrangements that permit only excited NH_3 molecules to interact with the radiation field. Another form of maser, based on three quantum levels, has been proposed by BLOEMBERGEN.[2] In this case a strong driving or pumping field is applied resonant with the outer pair of levels and is effective in so changing the equilibrium populations of the levels that an inverted susceptibility appears across one of the inner pair of levels. We shall calculate this susceptibility in the following sections, using an analysis that includes the effect of off-diagonal components of the density matrix. A similar approach has been used by ANDERSON[3] and JAVAN.[4] The results generally agree with those of BLOEMBERGEN, but they permit a discussion of certain questions not touched upon in his work. We are able to show, for instance, that the susceptibility offered by one pair of levels is usually unaffected by the cavity reaction on the other pair.* Another result is a saturation of the susceptibility that appears at high levels of the pumping field. This effect is also described by JAVAN.[4]

2. EQUATIONS OF MOTION

Let us consider the three-level quantum-mechanical system shown in Fig. 1, where the energy levels are E_1, E_2, and E_3 in order of ascending energy. We shall describe this system in terms

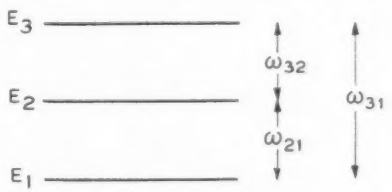


FIG. 1. Energy level scheme.

of its density matrix ρ_{nm}, where n and m run from one to three. The equations of motion for ρ_{nm},

* ANDERSON[3] describes a case in which the cavity reaction is important.

including relaxation terms, are

$$n = m$$

$$\dot{\rho}_{nn} = \sum_k (W_{kn}\rho_{kk} - W_{nk}\rho_{nn}) + \\ + \frac{1}{i\hbar}\sum_k (V_{nk}\rho_{kn} - \rho_{nk}V_{kn}) \tag{1}$$

$$n \neq m$$

$$\dot{\rho}_{nm} = -\frac{1}{\tau_{nm}}\rho_{nm} + \frac{1}{i\hbar}(E_n - E_m)\rho_{nm} + \\ + \frac{1}{i\hbar}\sum_k (V_{nk}\rho_{km} - \rho_{nk}V_{km}). \tag{2}$$

In these equations the terms W_{kn} are transition probabilities between the various levels and are subject to the restrictions of detailed balancing

$$0 = (W_{kn}\rho_{kk}^0 - W_{nk}\rho_{nn}^0) \tag{3}$$

where ρ_{nn}^0 represents the components of ρ_{nm} in thermal equilibrium. The quantities $\tau_{nm} = \tau_{mn}$ are relaxation times associated with the off-diagonal components of ρ_{nm}. The elements V_{nk} constitute a perturbation matrix existing because of the presence of electromagnetic fields. We suppose that there are three plane-polarized fields, separated in frequency, and interacting with the transitions $1 \to 3$, $1 \to 2$, and $2 \to 3$. The matrix components V_{nm} may then be written

$$V_{nm} = -\mu_{nm}H_{nm}\cos(\Omega_{nm}t - \phi_{nm}) \tag{4}$$

where μ_{nm} is the dipole moment associated with the transition $n \to m$. H_{nm}, Ω_{nm}, and ϕ_{nm} are the corresponding magnetic field, frequency, and phase angle. The positions of the indices in H_{nm} have no significance. We take Ω_{nm} to be positive if $n > m$ and place $\Omega_{mn} = -\Omega_{nm}$ and $\phi_{mn} = -\phi_{nm}$.

Let us now enter an interaction representation by making the substitution

$$\rho_{nm} = \sigma_{nm}\exp(-i\omega_{nm}t) \tag{5}$$

where

$$\omega_{nm} = \frac{1}{\hbar}(E_n - E_m).$$

Equations (1) and (2) become respectively

$$\dot{\sigma}_{nn} = \sum_k (W_{kn}\sigma_{kk} - W_{nk}\sigma_{nn}) - \\ - \sum_k \left\{\frac{H_{nk}}{2i\hbar}\right\}\Big[\mu_{nk}\sigma_{kn}\exp\{i[(\Omega_{kn} - \omega_{kn})t - \phi_{kn}]\} - \\ - \sigma_{nk}\mu_{kn}\exp\{-i[(\Omega_{kn} - \omega_{kn})t - \phi_{kn}]\}\Big] \tag{6}$$

$$\dot{\sigma}_{nm} = -\frac{1}{\tau_{nm}}\sigma_{nm} - \\ - \sum_k \left\{\frac{H_{nk}}{2i\hbar}\right\}\mu_{nk}\sigma_{km}\exp\{i[(\Omega_{kn} - \omega_{kn})t - \phi_{kn}]\} - \\ - \left\{\frac{H_{km}}{2i\hbar}\right\}\sigma_{nk}\mu_{km}\exp\{-i[(\Omega_{km} - \omega_{km})t - \phi_{km}]\}. \tag{7}$$

We next proceed by making the substitutions

$$A_{nk} = \frac{1}{2\hbar}H_{nk}\mu_{nk}\exp(i\phi_{nk}) \tag{8}$$

$$\sigma_{nk} = \lambda_{nk}\exp[-i(\Omega_{nk} - \omega_{nk})t] \tag{9}$$

and then obtain, keeping only the secular terms,

$$0 = \sum_k (W_{kn}\lambda_{kk} - W_{nk}\lambda_{nn}) + \\ + i\sum_k (A_{nk}\lambda_{kn} - \lambda_{nk}A_{kn}) \tag{10}$$

$$\left[-i(\Omega_{nm} - \omega_{nm}) + \frac{1}{\tau_{nm}}\right]\lambda_{nm} \\ = i\sum_k (A_{nk}\lambda_{km} - \lambda_{nk}A_{km}) \tag{11}$$

with the condition that

$$\Omega_{31} = \Omega_{32} + \Omega_{21}. \tag{12}$$

Expressions (10) and (11) represent six equations in the six independent components of the density matrix λ_{nm} and could be solved in general. However, we proceed in the next section under somewhat specialized conditions.

3. COMPONENTS OF THE DENSITY MATRIX

We shall assume that the diagonal elements of the density matrix are essentially controlled by a large radio-frequency magnetic field causing transitions between levels 1 and 3 and that this field is in resonance with these levels, so that

VOL. 4 1958

$\Omega_{31}=\omega_{31}$. Writing out explicitly four of the six equations contained in (10) and (11), we have

$$-(W_{12}+W_{13})\rho_{11}+W_{21}\rho_{22}+W_{31}\rho_{33}+ +i(A_{13}\lambda_{31}-\lambda_{13}A_{31})=0 \tag{13}$$

$$-(W_{21}+W_{23})\rho_{22}+W_{12}\rho_{11}+W_{32}\rho_{33}=0 \tag{14}$$

$$-(W_{31}+W_{32})\rho_{33}+W_{13}\rho_{11}+W_{23}\rho_{22}- -i(A_{13}\lambda_{31}-\lambda_{13}A_{31})=0 \tag{15}$$

$$\lambda_{13}=i\tau_{13}A_{13}(\rho_{33}-\rho_{11}) \tag{16}$$

since it is obvious that $\lambda_{nn}=\sigma_{nn}=\rho_{nn}$. Certain second-order terms have been dropped in these equations. Equation (15) is not independent of equations (13) and (14), but we may use the additional relation $\rho_{11}+\rho_{22}+\rho_{33}=1$, together with equations (13), (14) and (16), to obtain

$$(\rho_{33}-\rho_{22})=\frac{(\rho_{33}^0-\rho_{22}^0)+\frac{1}{\hbar^2}H_{13}^2|\mu_{13}|^2\tau_{13}T_e}{1+\frac{1}{\hbar^2}H_{13}^2|\mu_{13}|^2\tau_{13}T} \tag{17}$$

$$(\rho_{22}-\rho_{11})=\frac{(\rho_{22}^0-\rho_{11}^0)-\frac{1}{\hbar^2}H_{13}^2|\mu_{13}|^2\tau_{13}T_e}{1+\frac{1}{\hbar^2}H_{13}^2|\mu_{13}|^2\tau_{13}T} \tag{18}$$

$$(\rho_{33}-\rho_{11})=\frac{(\rho_{33}^0-\rho_{11}^0)}{1+\frac{1}{\hbar^2}H_{13}^2|\mu_{13}|^2\tau_{13}T}. \tag{19}$$

If a set of relaxation times T_{nm}, consistent with equation (3), are defined by

$$W_{nm}=\frac{1}{T_{nm}}\rho_{mm}^0, \tag{20}$$

then the quantities T and T_e are given by*

$$T=\frac{\frac{1}{T_{12}}\rho_{11}^0+\frac{1}{2}\left(\frac{1}{T_{12}}+\frac{1}{T_{23}}\right)\rho_{22}^0+\frac{1}{T_{23}}\rho_{33}^0}{\frac{1}{T_{13}T_{12}}\rho_{11}^0+\frac{1}{T_{12}T_{23}}\rho_{22}^0+\frac{1}{T_{13}T_{23}}\rho_{33}^0} \tag{21}$$

$$T_e=\frac{1}{2}\cdot\frac{\frac{1}{T_{12}}(\rho_{11}^0-\rho_{22}^0)+\frac{1}{T_{23}}(\rho_{33}^0-\rho_{22}^0)}{\frac{1}{T_{13}T_{12}}\rho_{11}^0+\frac{1}{T_{12}T_{23}}\rho_{22}^0+\frac{1}{T_{13}T_{23}}\rho_{33}^0}. \tag{22}$$

If the remaining two equations contained in (11) are written out explicitly, we have

$$\left[i(\Omega_{32}-\omega_{32})+\frac{1}{\tau_{21}}\right]\lambda_{21}+iA_{31}\lambda_{23} =iA_{21}(\rho_{11}-\rho_{22})+iA_{23}\lambda_{31} \tag{23}$$

$$\left[i(\Omega_{32}-\omega_{32})+\frac{1}{\tau_{23}}\right]\lambda_{23}+iA_{13}\lambda_{21} =iA_{23}(\rho_{33}-\rho_{22})+iA_{21}\lambda_{13} \tag{24}$$

where we have set $(\Omega_{21}-\omega_{21})=-(\Omega_{32}-\omega_{32})$ as follows from the assumption $\Omega_{31}=\omega_{31}$. These are the basic equations of the three-level maser. We note first that the free oscillation of the off-diagonal components of the density matrix are coupled in the presence of the pumping field H_{13}. The coupling clearly arises from the fact that ρ_{23} can beat with H_{13} to produce a frequency that excites ρ_{21}, and similarly ρ_{21} can beat with H_{13} to excite ρ_{23}. Next we see that this set of coupled equations is driven by the signal fields H_{12} and H_{23} in a way depending on the population difference of the corresponding levels. Finally, the equations are driven by terms arising from a beat between the signal fields and the oscillation of ρ_{13}.

The quantities $(\rho_{11}-\rho_{22})$, $(\rho_{33}-\rho_{22})$, and λ_{13} have already been found. We may proceed then to solve equations (23) and (24) simultaneously, with the result

$$\lambda_{21}=\frac{i\left[i(\Omega_{32}-\omega_{32})+\frac{1}{\tau_{23}}\right][A_{21}(\rho_{11}-\rho_{22})+A_{23}\lambda_{31}]+A_{31}[A_{23}(\rho_{33}-\rho_{22})+A_{21}\lambda_{13}]}{\left[i(\Omega_{32}-\omega_{32})+\frac{1}{\tau_{21}}\right]\left[i(\Omega_{32}-\omega_{32})+\frac{1}{\tau_{23}}\right]+|A_{13}|^2} \tag{25}$$

* Definition of an effective time T_e has also been considered by LLOYD and PAKE.[6]

$$\lambda_{23}=\frac{i\left[i(\Omega_{32}-\omega_{32})+\frac{1}{\tau_{12}}\right][A_{23}(\rho_{33}-\rho_{22})+A_{21}\lambda_{13}]+A_{13}[A_{21}(\rho_{11}-\rho_{22})+A_{22}\lambda_{31}]}{\left[i(\Omega_{32}-\omega_{32})+\frac{1}{\tau_{21}}\right]\left[i(\Omega_{32}-\omega_{32})+\frac{1}{\tau_{23}}\right]+|A_{13}|^2}. \tag{26}$$

We have now obtained all components of the density matrix and can proceed to evaluate the susceptibilities.

4. SUSCEPTIBILITIES

Let us now characterize the physical situation as follows. The paramagnetic sample is in a resonant cavity in the presence of three radio-frequency fields. One of these is the pumping field H_{13}. The second field, H_{23}, will be a signal field introduced into the cavity from an external source. The response of the sample to the signal field will produce an oscillating dipole moment M_{12} between levels 1 and 2. This dipole will create a third field, H_{12}, in the cavity in a way determined by the cavity "susceptibility" $\chi_c=\chi_c'-i\chi_c''$. That is

$$M_{12}=\chi_c'H_{12}\cos(\Omega_{21}\tau-\phi_{21})+ +\chi_c''H_{12}\sin(\Omega_{21}\tau-\phi_{21}) \tag{27}$$

where χ_c can be shown in a standard way to be given by

$$\frac{1}{\chi_c}=\sum_n\frac{4\pi F_n[\omega^2-i(\omega\omega_n/Q_n)]}{(\omega_n{}^2-\omega^2)+i(\omega\omega_n/Q_n)} \simeq\sum_n\frac{4\pi F_n\omega}{2(\omega_n-\omega)+i(\omega_n/Q_n)}. \tag{28}$$

Here the sum is carried over all modes of the cavity, with ω_n, Q_n, and F_n (the filling factor) being characteristics of a particular mode. Expressed in terms of the density matrix, M_{12} is given by

$$M_{12}=N(\rho_{12}\mu_{21}+\rho_{21}\mu_{12}) =N[\lambda_{12}\exp(i\Omega_{21}\tau)\mu_{21}+\lambda_{21}\exp(-i\Omega_{21}\tau)\mu_{12}] \tag{29}$$

where N is the number of spins per unit volume. From equations (27) and (29) we find with a little manipulation

$$(\chi_c'+i\chi_c'')A_{21}=\frac{N|\mu_{12}|^2}{\hbar}\lambda_{21}. \tag{30}$$

Similarly, if the susceptibility of the sample offered to the signal field by transitions between levels 2 and 3 is $\chi=\chi'-i\chi''$, we find

$$(\chi'-i\chi'')A_{23}=\frac{N|\mu_{23}|^2}{\hbar}\lambda_{23}. \tag{31}$$

If λ_{21} and λ_{23} are substituted from equations (25) and (26) into equations (30) and 31), there are obtained two homogeneous equations in A_{21} and A_{23}. Placing the determinant of the coefficients equal to zero, one obtains an equation for the susceptibility χ

$$\chi=\chi_{32}\left[i\left(\frac{d_{21}+rb}{d_{21}d_{32}+b}\right)+\frac{b\left(\frac{1+Rd_{21}}{d_{21}d_{32}+b}\right)\left(\frac{1+rd_{32}}{d_{21}d_{32}+b}\right)\chi_{12}}{\chi_c{}^*-i\left(\frac{d_{32}-Rb}{d_{21}d_{32}+b}\right)\chi_{12}}\right] \tag{32}$$

where we make the following definitions:

$$d_{21}=1-i(\Omega_{21}-\omega_{21})\tau_{12}$$

$$d_{32}=1+i(\Omega_{32}-\omega_{32})\tau_{23}$$

$$R=\frac{\tau_{13}(\rho_{11}-\rho_{33})}{\tau_{12}(\rho_{11}-\rho_{22})},\qquad r=\frac{\tau_{13}(\rho_{11}-\rho_{33})}{\tau_{23}(\rho_{33}-\rho_{22})}$$

$$b=\frac{1}{4\hbar^2}H_{13}{}^2|\mu_{13}|^2\tau_{12}\tau_{23}$$

$$\chi_{12}=\frac{1}{\hbar}N|\mu_{12}|^2\tau_{12}(\rho_{11}-\rho_{22})$$

VOL. 4 1958

$$\chi_{32} = \frac{1}{\hbar} N|\mu_{23}|^2 \tau_{23}(\rho_{33} - \rho_{22}).$$

Equation (32) is the most general expression for the susceptibility of a three-level maser. The susceptibility, together with the characteristics of the microwave cavity at the frequency Ω_{32}, determine all the important properties of the maser except its noise behavior.

5. APPROXIMATE EQUATION FOR THE SUSCEPTIBILITY

We shall next consider a special set of conditions, appropriate to the operation of a three-level maser, which allows equation (32) to be considerably simplified. Suppose first that the various energy differences are small enough compared with the operating temperature so that one may write approximately

$$\left.\begin{aligned} \rho_{11}^0 - \rho_{22}^0 &= \frac{E_2 - E_1}{3kT} \\ \rho_{11}^0 - \rho_{33}^0 &= \frac{E_3 - E_1}{3kT} \\ \rho_{22}^0 - \rho_{33}^0 &= \frac{E_3 - E_2}{3kT}. \end{aligned}\right\} \tag{33}$$

Let us also assume that $T_{12} \ll T_{13}$ or T_{23}, a circumstance that can be realized in the manner reported by SCOVIL *et al.*[6] In that case, equation (18) indicates that very closely

$$(\rho_{11} - \rho_{22}) = (\rho_{11}^0 - \rho_{22}^0). \tag{34}$$

From equation (19) and the definition of b, we find

$$(\rho_{11} - \rho_{33}) = \frac{(\rho_{11}^0 - \rho_{33}^0)}{1 + 4b(T/\tau)} \tag{35}$$

where we have assumed all the off-diagonal relaxation times equal to τ. From equations (34) and (35) there is found

$$R = \frac{E_3 - E_1}{E_2 - E_1} \cdot \frac{1}{1 + \dfrac{4T}{\tau b}}. \tag{36}$$

It is also easily shown that

$$r = \frac{R}{1 - R} \tag{37}$$

and

$$(\rho_{33} - \rho_{22}) = (\rho_{11}^0 - \rho_{22}^0)(1 - R). \tag{38}$$

A slight rearrangement of equation (32) gives

$$\chi = i\chi_{32}\left(\frac{d_{21} + rb}{d_{21}d_{32} + b}\right) \times \times \left[1 + \frac{b\left(\dfrac{1 + Rd_{21}}{d_{21}d_{32} + b}\right)\left(\dfrac{1 + rd_{32}}{d_{21} + rb}\right)}{\dfrac{i\chi_c^*}{\chi_{21}} + \left(\dfrac{d_{32} - Rb}{d_{21}d_{32} + b}\right)}\right]. \tag{39}$$

We now subject the second term in brackets in equation (39) to a close examination based on the following considerations:

(*a*) We shall be interested in the susceptibility χ only for Ω_{32} fairly close to ω_{32} and for Ω_{21} correspondingly close to ω_{21}.

(*b*) The cavity susceptibility χ_c will be controlled by the cavity mode whose resonance is closest to Ω_{21} and is found from equation (28) to be

$$\chi_c = \frac{2(\omega_0 - \Omega_{21})}{4\pi F\omega_0} + i\frac{1}{4\pi FQ}. \tag{40}$$

We assume typical values of $F = \frac{1}{10}$ and $Q = 5000$. If $\omega_0 \simeq \Omega_{21}$, we shall say that the resonance is close. There must always be a cavity resonance near ω_{32} in any useful maser. If there is no resonance closer to Ω_{21}, and if the level separations are such that $|\Omega_{21} - \omega_0| > \omega_0$, we shall say that the resonance is remote.

(*c*) We assume that the energy difference $(E_2 - E_1)$ corresponds to a frequency in the neighborhood of 24,000 Mc/s and that the working temperature is a few degrees absolute. Then $(\rho_{11}^0 - \rho_{22}^0) \simeq \frac{1}{10}$. If the spin density is about 10^{20} and τ about 10^{-8} sec, we find $\chi_{12} \simeq 10^{-2}$.

The detailed analysis allows one to conclude that the second term may be neglected under the following conditions:

(*a*) If ω_0 is remote and $1 - R > 0$

(*b*) If ω_0 is close, $1 - R > \frac{1}{10}$, and

$$\frac{\tau}{T} < 4\left(\frac{E_2 - E_1}{E_3 - E_2}\right)10^{-4}.$$

The condition on R avoids some complications in

the behavior of χ at low levels of the pumping field, but leaves available the interesting range. In essence, the second term, may be neglected if ω_0 is remote, and may still be neglected if ω_0 is close, provided τ/T is small enough. With these restrictions, we write for the susceptibility

$$\chi = i\chi_{12}\frac{(1-R)(1+i\Delta)Rb}{(1+i\Delta)^2+b} \tag{41}$$

where we have placed $\Delta = (\Omega_{32}-\omega_{32})\tau = -(\Omega_{21}-\omega_{21})\tau$ and assumed $|\mu_{23}|^2 = |\mu_{12}|^2$. In this approximation, χ no longer depends upon the cavity susceptibility. The real and imaginary parts of $\chi = \chi'-i\chi''$ may be easily found from equation (41) to be

$$\frac{\chi'}{\chi_{12}} = \frac{\Delta[(1-R)+(3R-1)b+(1-R)\Delta^2]}{(1+b-\Delta^2)^2+4\Delta^2} \tag{42}$$

$$\frac{\chi''}{\chi_{12}} = \frac{(1+b)(1-R+Rb)+(1-R-Rb)\Delta^2}{(1+b-\Delta^2)^2+4\Delta^2}. \tag{43}$$

If b is set equal to zero in equations (41), (42), and (43), the susceptibilities then correspond to BLOEMBERGEN's original results.

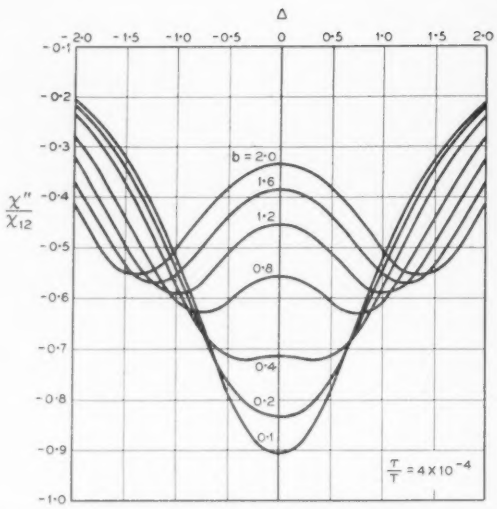


FIG. 3. Imaginary part of the normalized susceptibility as a function of $\Delta = (\Omega_{32}-\omega_{32})\tau$, for $\tau/T = 4\times10^{-4}$ and various values of $b = |(1/2\hbar)H_{13}\mu_{13}\tau|^2$.

In Figs. 2–6 we have plotted a series of susceptibility curves calculated from equations (42) and (43). These curves show the dependence of χ'/χ_{12} and of χ''/χ_{12} upon $\Delta = (\Omega_{32}-\omega_{32})\tau$ and are plotted

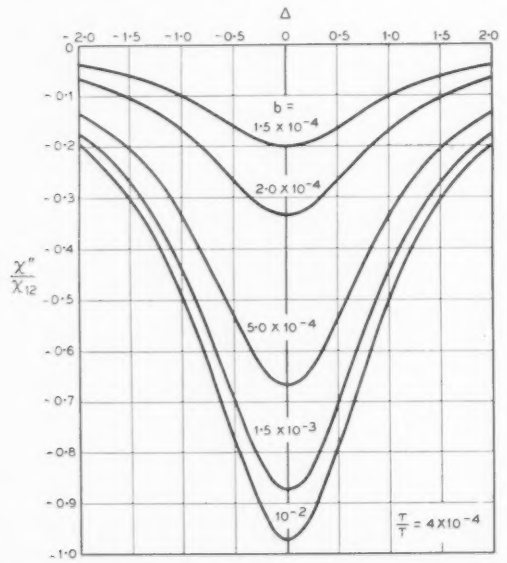


FIG. 2. Imaginary part of the normalized susceptibility as a function of $\Delta = (\Omega_{32}-\omega_{32})\tau$, for $\tau/T = 4\times10^{-4}$ and various values of $b = |(1/2\hbar)H_{13}\mu_{13}\tau|^2$.

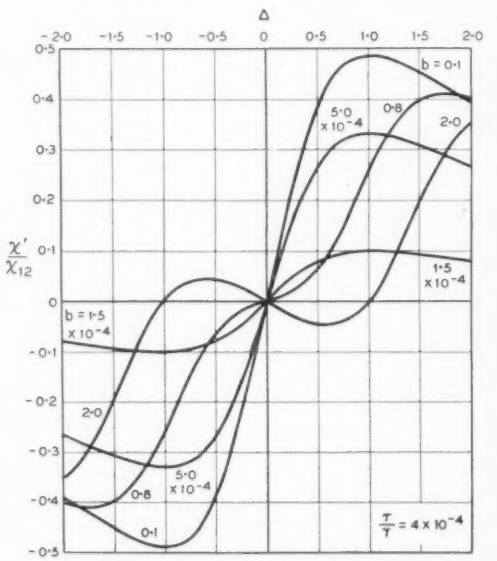


FIG. 4. Real part of the normalized susceptibility as a function of $\Delta = (\Omega_{32}-\omega_{32})\tau$, for $\tau/T = 4\times10^{-4}$ and various values of $b = |(1/2\hbar)H_{13}\mu_{13}\tau|^2$.

VOL. 4 1958

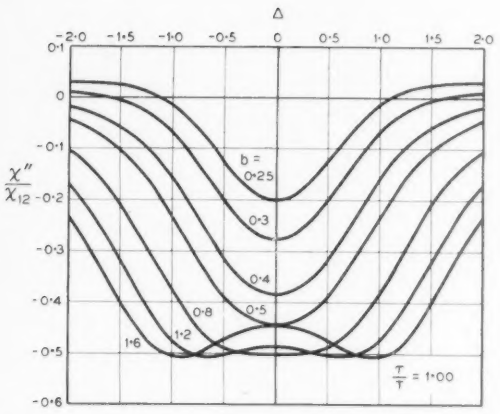


FIG. 5. Imaginary part of the normalized susceptibility as a function of $\Delta = (\Omega_{42} - \omega_{32})\tau$, for $\tau/T = 1{\cdot}00$ and various values of $b = |(1/2\hbar)H_{13}\mu_{13}\tau|^2$.

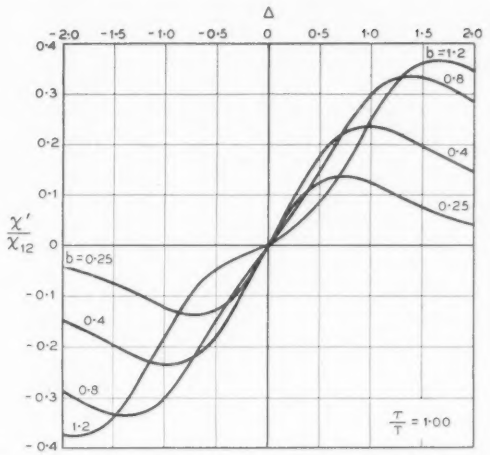


FIG. 6. Real part of the normalized susceptibility as a function of $\Delta = (\Omega_{32} - \omega_{32})\tau$, for $\tau/T = 1{\cdot}00$ and various values of $b = |(1/2\hbar)H_{13}\mu_{13}\tau|^2$.

for various values of $b = |\tau/2\hbar H_{13}\mu_{13}|^2$. Two extreme values have been chosen for the ratio of off-diagonal to diagonal relaxation times. In Figs. 2, 3, and 4, we have chosen $\tau/T = 4\times10^{-4}$; the condition of validity for equation (41) is then met for ω_0 either close or remote provided b is greater than $1{\cdot}2\times10^{-4}$. Figs. 2 and 3 show the imaginary part of the susceptibility, while the real part is shown in Fig. 4. In Figs. 5 and 6, we take $\tau/T = 1$; the curves are then valid only for ω_0 remote and $b > \frac{1}{4}$. The imaginary part of the susceptibility is presented in Fig. 5 and the real part in Fig. 6.

These curves exhibit the increasing negative value of χ'' as the pumping signal specified by b is increased. The most striking feature of the curves is the splitting of the line into a doublet when b becomes sufficiently large. This behavior has been discussed by JAVAN[4] and is illustrated in Fig. 2 of his paper. The effect has also been discussed by AUTLER and TOWNES[7] in connection with the molecule OCS. The splitting of the line is necessarily accompanied by an anomalous behavior of the real part of the susceptibility as may be seen in Figs. 4 and 6.

Acknowledgements—The author would like to acknowledge many helpful discussions with P. W. ANDERSON and H. SUHL. In particular he would like to thank A. JAVAN for the opportunity of examining his paper prior to its publication.

REFERENCES

1. GORDON J. P. ZEIGER H. J. and TOWNES C. H. *Phys. Rev.* **99,** 1264 (1955).
2. BLOEMBERGEN N. *Phys. Rev.* **104,** 324 (1956).
3. ANDERSON P. W. *J. Appl. Phys.* **28,** 1049, (1957).
4. JAVAN A. *Phys. Rev.* To be published.
5. SCOVIL H. E. D., FEHER G. and SEIDEL H. *Phys. Rev.* **105,** 762 (1957).
6. LLOYD J. P. and PAKE G. E. *Phys. Rev.* **94,** 579 (1954).
7. AUTLER S. H. and TOWNES C. H. *Phys. Rev.* **100,** 703 (1955).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 278–282.

QUANTUM ANALOG OF THE FERROMAGNETIC MICROWAVE AMPLIFIER

H. SUHL

Bell Telephone Laboratories, Inc., Murray Hill, N. J.

(*Received* 28 *August* 1957)

Abstract—Certain resemblances between the modes of operation of the three-level maser and the ferromagnetic microwave amplifier are shown to be superficial by constructing a quantum-mechanical model of the latter device. It is shown that while establishment of a negative temperature for two levels is essential in the three-level maser, it is inessential in the analog of the ferromagnetic amplifier, which in essence depends only on the time-varying part of the density matrix.

VOL.
4
1958

INTRODUCTION

THERE have been two recent additions to the family of solid-state microwave devices: The three-level paramagnetic maser proposed by BLOEMBERGEN[1] and put into practice by SCOVIL et al.,[2] and the ferromagnetic amplifier proposed by the present author[3,4] and put into practice by WEISS.[5]

Superficially, the operating conditions of the two devices are quite similar. Both require an oscillatory power supply (the "pump"), of a frequency higher than that of the signal to be amplified. In both, the material should be approximately resonant to the frequency of the supply; both devices must have two additional resonances, the sum of whose frequencies equals the pump frequency. Yet the two operate on entirely different principles. A direct comparison is not immediately possible, since the simplest description of the BLOEMBERGEN device is quantum-mechanical, while the ferromagnetic amplifier is most naturally analyzed in classical terms.[4] In this paper, we discuss a quantum model of the latter device, in order to bring the differences into focus. While this model has little in common with a ferromagnet as such, the analysis of its operation as an amplifier exactly parallels that of the ferromagnetic device.

To begin with, we note that the only function of a third level in the three-level maser as originally proposed is to make possible the establishment of an inverted population in two of the levels. Thus, if the material has three levels, numbered in order of increasing energy $E_i (i = 1, 2, 3)$, the "pump", working at a frequency $(E_3 - E_1)/\hbar$, raises the average population in level 3 above that of level 2. Amplification by stimulated emission at frequency $(E_3 - E_2)/\hbar$ can then occur under suitable conditions.[1] Without doing justice to the many subtleties involved in a detailed discussion of the original three-level maser, we may state the role of the three frequencies thus: the signal at the pump frequency serves to reverse the normal Boltzmann populations of levels 2 and 3; the frequency $(E_3 - E_2)/\hbar$ is the one to be amplified, and the difference frequency $(E_2 - E_1)/\hbar$ serves no special purpose at all. In establishing the condition for amplification, all one needs to know about the statistics of the material is the average populations of level 2 and 3. That is to say, even though the three-level maser is not in thermal equilibrium, it is possible to describe its simpler aspects (but not all aspects) in terms of the diagonal elements of the density matrix alone.

The ferromagnetic amplifier, on the other hand, cannot be treated on the basis of average-level populations alone. Instead, the time-dependent off-diagonal, as well as the fluctuating parts of the on-diagonal elements of the density matrix are required. The average-level populations can be normal or inverse, and their precise values can affect only the efficiency, not the principle, of operation.

The model we shall consider will be described in terms of the density matrix; therefore we list briefly the papers available to date which treat the three-level maser in similar terms:

(i) A paper by ANDERSON.[6] Besides giving the condition for maser action of the original BLOEMBERGEN proposal (inversion of the normal population in level 2 and 3), ANDERSON suggests some further devices some of which are "hybrids" between three-level masers and the ferromagnetic amplifier.

(ii) A paper by CLOGSTON.[7] This exhibits the detailed behavior of the three-level maser over a very wide range of pumping powers. Rather than look for the onset of instability, CLOGSTON determines the pumping power at which the imaginary part of the sample susceptance equals that of the circuit. The result is, of course, the same: reversal of populations between level 2 and 3 is required.

(iii) A paper by JAVAN.[8] While this does not mention the density matrix directly, it does follow the motion of the amplitudes of wave functions and of time integrals of products of these amplitudes. Presumably it is therefore equivalent to the density-matrix method, and the results are in fact the same as those of CLOGSTON,[7] in so far as they refer to the BLOEMBERGEN Maser.

THE MODEL

The model we shall consider may be described as the paramagnetic analog of the ferromagnetic amplifier in so-called "electromagnetic operation."[3,4] A sample of N non-interacting spins of $\frac{1}{2}$ is placed in a steady field. The assembly is in contact with a heat reservoir, tending to keep the population of the upper and lower Zeeman levels at their appropriate Boltzmann values. The sample is surrounded by a microwave cavity so constructed as to resonate at the two frequencies, ω_1, ω_2, whose sum equals ω, the Zeeman frequency. The two cavity modes must be such that the r.f. magnetic field of at least one has components along the direction of the d.c. field (the Oz axis, say), while the r.f. magnetic field of the other has a component in a plane normal to Oz. In addition, the microwave structure has provisions for applying the "pump" at the Zeeman frequency with magnetic field in the transverse plane.

We will show that, as in the case of the ferromagnetic amplifier, oscillations can set in at both frequencies when the pumping field exceeds a certain threshold value. This is, of course, not too surprising, since in this simple system the quantum-mechanical equations of motion of each spin are identical in form with the corresponding classical equation of motion for the magnetization of a ferromagnet. However, for direct comparison with maser theory, which is always kept in terms of level population, it is better to analyze the present model in terms of the density matrix.

Since the spins are mutually independent, it is sufficient to consider the motion of the two-by-two density matrix ρ of a single spin in a representation in which S_z is diagonal. To account for the presence of dissipation within the material we introduce the relaxation times T_1 and T_2 for the on- and off-diagonal element of ρ. If V is the total perturbing energy due to the field of the pump and to the fields in the two modes of the cavity, the three equations of motion for the elements of ρ may be written:

$$\left.\begin{aligned}
&\text{(a)}\quad i\hbar\dot{\rho}_{11} = V_{12}\rho_{21}-\rho_{12}V_{21}-\frac{i\hbar(\rho_{11}-\rho_{11}^{0})}{T_1}\\
&\text{(b)}\quad i\hbar\dot{\rho}_{22} = V_{21}\rho_{12}-\rho_{21}V_{12}-\frac{i\hbar(\rho_{22}-\rho_{22}^{0})}{T_1}\\
&\text{(c)}\quad i\hbar\dot{\rho}_{21} = \hbar\omega\rho_{21}\rho_{21}+V_{21}(\rho_{11}-\rho_{22})+\\
&\qquad\qquad +(V_{22}-V_{11})\rho_{21}-\frac{i\hbar\rho_{21}}{T_2}
\end{aligned}\right\}\quad \text{(i)}$$

where ρ_{11}^{0}, ρ_{22}^{0} are the Boltzmann factors for thermal equilibrium. From (i) (a, b), we get

$$i\hbar\dot{\Delta} = 2(V_{12}\rho_{21}-\rho_{12}V_{21})-\frac{i\hbar}{T_1}(\Delta-\Delta^{0}) \qquad \text{(ii)}$$

where $\Delta = \rho_{11}-\rho_{22}$, $\Delta^{0} = \rho_{11}^{0}-\rho_{22}^{0}$.

The two-cavity modes have h fields that will be denoted by $\frac{1}{2}\boldsymbol{h}_j[b_j\exp(i\omega_j t)+b_j^{*}\exp(-i\omega_j t)]$ $(j=1,2)$, where the $\boldsymbol{h}_j$ are real vectors and the b_j are coefficients varying slowly with time. The pump-field will simply be written

$$\tfrac{1}{2}\boldsymbol{h}_p[\exp(i\omega t)+\exp(-i\omega t)].$$

Clearly

$$V_{21} = -\frac{g\mu_B}{2}\sum_{j=1}^{2}(\boldsymbol{s}_{21}\cdot\boldsymbol{h}_j)[b_j\exp(i\omega_j t)+b_j^*\exp(-i\omega_j t)]-\frac{g\mu_B}{2}(\boldsymbol{s}_{21}\cdot\boldsymbol{h}_p)\times$$
$$\times[\exp(i\omega t)+\exp(-i\omega t)]$$
$$\equiv\sum_{j=1}^{2}W_j[b_j\exp(i\omega_j t)+b_j^*\exp(-i\omega_j t)]+W_p[\exp(i\omega t)+\exp(-i\omega t)],\ \text{say.}$$
$$V_{11} = -\frac{g\mu_B}{2}\cdot\frac{1}{2}\sum_{j=1}^{2}h_{jz}[b_j\exp(i\omega_j t)+b_j^*\exp(-i\omega_j t)]$$
$$= -V_{22}$$
$$\equiv\sum_{j=1}^{2}U_j[b_j\exp(i\omega_j t)+b_j^*\exp(i\omega_j t)],\ \text{say} \qquad \text{(iii)}$$

where μ_B is the Bohr magneton and g the g-factor of the spins. We attempt a solution of equation (i)(c) by writing

$$\rho_{21} = A\exp(i\omega_1 t)+B\exp(-i\omega_1 t)+$$
$$+C\exp(i\omega_2 t)+D\exp(-i\omega_2 t)+E\exp(-i\omega t) \qquad \text{I(iv)}$$

where we suppose that, owing to the presence of the pump, A, B, C, and D will be slowly varying functions of time, rather than constants. The presence of E is due chiefly to the pump and may be calculated (for small excitations of the modes) by substituting (iii), (iv) in (i), equating to zero all signal terms like A, B, C, D, W_1, W_2, and retaining only the pump. Neglecting all non-secular terms and setting $\Delta = \Delta^0$, we get:

$$E = \frac{1}{i\hbar\nu_2}W_p\Delta^0; \qquad (\nu_2 = 1/T_2) \qquad \text{(v)}$$

Next, to determine A, B, C, D, equation (iv) is substituted in (i)c, and coefficients of terms with the same frequency are equated, giving

$$\hbar\left[i\left(\nu_2+\frac{d}{dt}\right)-(\omega_1+\omega)\right]A = W_1 b_1\Delta^0$$
$$\hbar\left[i\left(\nu_2+\frac{d}{dt}\right)+(\omega_1-\omega)\right]B = W_1 b_1^*\Delta^0-2U_2 b_2 E \qquad \text{(vi)}$$

$$\hbar\left[i\left(\nu_2+\frac{d}{dt}\right)-(\omega_2+\omega)\right]C = W_2 b_2\Delta^0$$
$$\hbar\left[i\left(\nu_2+\frac{d}{dt}\right)+(\omega_2-\omega)\right]D = W_2 b_2^*\Delta^0-2U_2 b_1 E \qquad \text{(vi)}$$

Since we are dealing with a type of operation in which neither ω_1 nor ω_2 is equal to ω, (ν_2+d/dt) may be neglected on the left-hand side of these equations. In relating the field amplitudes to the expectation value $\langle\mathbf{M}\rangle$ through Maxwell's equations, we shall also require the time-varying part of $\langle M_z\rangle$. Since this is proportional to Δ, we also require the time-varying part, $\tilde{\Delta}$, of Δ, even though we consider its time-independent part to be practically equal to Δ^0. The same kind of analysis that led to equation (vi) gives

$$\tilde{\Delta} = R\exp(i\omega_1 t)+R^*\exp(-i\omega_1 t)+$$
$$+Q\exp(i\omega_2 t)+Q^*\exp(-i\omega_2 t)$$

where, approximately

$$R = \frac{2E^*W_2 b_2^*}{\hbar\Omega_1}+\text{small terms, of order}\frac{\nu_2}{\omega_{1,2}}$$
$$Q = \frac{2E^*W_1 b_1^*}{\hbar\Omega_2}+\text{small terms, of order}\frac{\nu_2}{\omega_{1,2}} \qquad \text{(vii)}$$

VOL. 4 1958

Thus we have determined how the cavity fields drive ρ, and hence how they drive the magnetization. The equations governing the variation of the cavity fields are now determined by requiring self-consistency, i.e. that the magnetization should drive the cavity fields according to Maxwell's equations. (We treat the electromagnetic fields classically.) Assuming the sample to be sufficiently small, we may use perturbation theory as follows: the two modes $\boldsymbol{h}_1, \boldsymbol{h}_2$ and the associated electric fields $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ may be chosen to satisfy

$$\nabla \times \boldsymbol{h}_j = \frac{\omega_j}{c}\boldsymbol{e}_j; \qquad \nabla \times \boldsymbol{e}_j = \frac{\omega_j}{c}\boldsymbol{h}_j$$

$$\nabla . \boldsymbol{h}_j = \nabla . \boldsymbol{e}_j = 0 \; (j=1,2)$$

where the $\boldsymbol{e}_j$ are normal to the cavity walls. It can be shown that the two modes are orthogonal, i.e. that

$$\int_{\text{cavity}} \boldsymbol{h}_1 . \boldsymbol{h}_2 \, dv = \int_{\text{cavity}} \boldsymbol{e}_1 . \boldsymbol{e}_2 \, dv = 0.$$

Substituting the total magnetic field

$$\boldsymbol{h} = \sum_{j=1}^{2} [b_i \exp(i\omega_j t) + c.c.]\boldsymbol{h}_j,$$

and the total electric field

$$\boldsymbol{k} = \Sigma \, [a_j \exp(i\omega t) + c.c.]\boldsymbol{e}_j,$$

into Maxwell's equations

$$\nabla \times \boldsymbol{h} = \frac{1}{c}\dot{\boldsymbol{e}}; \quad \nabla \times \boldsymbol{e} = -\frac{1}{c}\dot{\boldsymbol{h}} - \frac{1}{c}\langle \dot{\boldsymbol{M}} \rangle,$$

and using the orthogonality relations to eliminate a_j, we find

$$\frac{d^2 b_j}{dt^2} + 2i\omega_j \frac{db_j}{dt} = -4\pi . \text{ coeff. of } \exp(i\omega_j t) \text{ in } \frac{d^2}{dt^2} \int_{\text{sample}} \langle \boldsymbol{M} \rangle . \boldsymbol{h}_j \, dv \int_{\text{cavity}} \boldsymbol{h}_j{}^2 \, dv$$

Here $\langle \mathbf{M} \rangle = g\mu_B n_0 \operatorname{Tr} \rho \mathbf{s}$, the expectation value of the magnetization, where n_0 is the number of spins per unit volume. Since the b_j are assumed to vary slowly, we have $\omega_j(db_j/dt) \gg (d^2 b_j/dt^2)$, so that the second derivative on the left-hand side may be neglected. Similarly the differentiations of the right-hand side will produce terms of order $\omega_j{}^2 b_j$, $\omega_j b_j$, $d^2 b_j/dt^2$, and for the same reason as above the first of these is important, so that d^2/dt^2 may be replaced by $-\omega_j{}^2$ on the right-hand side. Finally, we must take account of dissipation in the cavity. It does not matter much how this is done; in the present instance it is best accomplished by adding a term $(i\omega_j{}^2/Q_j)b_j$ to the left-hand side, where Q_j is the quality factor of the cavity in mode j. Such a term evidently leads to the correct free decay of $b_j[\sim \exp(-\omega_j t/2Q_j)]$ if b_j is neglected. Dividing through by $i\omega_j$, we finally obtain

$$\frac{db_j}{dt} + \frac{\omega_j}{2Q_j} b_j = -4\pi i\omega_j . \text{ coeff. of } \exp(i\omega_j t) \text{ in } \int_{\text{sample}} \langle \boldsymbol{M} \rangle . \boldsymbol{h}_j . dv \Big/ \int_{\text{cavity}} \boldsymbol{h}_j{}^2 \, dv \tag{viii}$$

Now

$$\langle \boldsymbol{M} \rangle = g\mu_B n_0 \operatorname{Tr}(\rho \boldsymbol{s}) = g\mu_B n_0 [\rho_{21}\boldsymbol{s}_{12} + \rho_{12}\boldsymbol{s}_{21} + \rho_{22}\boldsymbol{s}_{22} + \rho_{11}\boldsymbol{s}_{11}]$$

Hence, and since $\mathbf{s}_{12}$ has only transverse components,

$$\langle \boldsymbol{M}(\omega_1)_{\text{transverse}} \rangle = -\frac{g\mu_B n_0 \Delta^0}{\hbar} \left[\frac{W_1 \boldsymbol{s}_{12}}{\omega_1 + \omega} + \frac{W_1{}^* \boldsymbol{s}_{21}}{\omega_1 - \omega} \right] b_1 - \frac{2\Delta^0 g\mu_B n_0 U_2 W_p{}^* \boldsymbol{s}_{21} b_2{}^*}{i(\hbar\omega_2)(\hbar\nu_2)} \tag{ix}$$

$$\langle M_z(\omega_1) \rangle = -g\mu_B n_0 \Delta^0 \frac{W_2 W_p}{\hbar\omega_1 \hbar\nu_2} b_2{}^* \tag{x}$$

and similarly for the components at frequency ω_2. When equation (ix) is substituted in (viii), it is seen that the first bracket on the right of (viii) gives only a small non-resonant frequency shift. Neglecting such terms, and restoring all the abbreviations, we find

$$\left(\frac{d}{dt} + \frac{1}{2Q_1} \right) b_1 = \frac{i\omega_1 \pi \Delta^0 g^2 \mu_B{}^2 W_p}{\hbar\nu_2 \int \boldsymbol{h}_1{}^2 \, dv} \times \left[\frac{1}{\hbar\omega_2} \int (h_{2z} s_z)(s_{21}{}^+ h_1{}^-) \, dv + \frac{1}{\hbar\omega_1} \int (h_{1z} s_z)(s_{21}{}^+ h_2{}^-) \, dv \right] b_2{}^* = \kappa_{12} b_2{}^*, \text{ say.} \tag{xi}$$

The trial solution $b_1 = B_1 \exp \lambda t$, $b_2 = B_2 \exp \lambda t$ will satisfy equation (xi) and the "adjoint" equation (with 1 or 2 interchanged and complex conjugates taken) if

$$\lambda^2 + \frac{\omega}{2}\left(\frac{1}{Q_1} + \frac{1}{Q_2}\right)\lambda + \frac{\omega_1\omega_2}{4Q_1Q_2} - \kappa_{12}\kappa_{21}^* = 0 \tag{xii}$$

It is easily seen that $\kappa_{12}\kappa_{21}^*$ is real and positive. One of the roots of equation (xii) is therefore positive when

$$\kappa_{12}\kappa_{21}^* > \frac{\omega_1\omega_2}{4Q_1Q_2}$$

and then oscillation in both modes will build up. This condition is easily transcribed into a threshold for the necessary "pumping" field. Just below the threshold, amplification will take place at either frequency, as well as conversion, with gain, from one frequency to the other.

We note that κ_{12} contains Δ^0, the *equilibrium* population difference between lower and upper state. This is because, in order to emphasize our point, we replaced the actual difference which should enter κ_{12} by the equilibrium value. Actually the sign of the difference is unimportant, since the instability criterion depends on the square of the difference. To minimize the threshold power, the population difference should be as large as possible, whatever its sign.

CONCLUSIONS

We have discussed a quantum analog of the ferromagnetic amplifier (in electromagnetic operation). We selected a paramagnetic two-level system, and found that it could act as an amplifier, whether the upper state had a greater or smaller population than the ground state. The average population playing a secondary role, its description hinges solely on the coupling together of different frequency components of the density matrix by the action of the pump.

VOL. 4 1958

REFERENCES

1. Bloembergen N. *Phys. Rev.* **104,** 324 (1956).
2. Scovil H. E. D., Feher G. and Seidel H. *Phys. Rev.* **105,** 762 (1957).
3. Suhl H. *Phys. Rev.* **106,** 384 (1957).
4. Suhl H. *J. Appl. Phys.* **28,** 1225 (1957).
5. Weiss M. T. *Phys. Rev.* **107,** 317 (1957).
6. Anderson P. W. *J. Appl. Phys.* **28,** 1049 (1957).
7. Clogston A. M. *J. Phys. Chem. Solids* **4,** 271 (1958).
8. Javan A. *Phys. Rev.* **107,** 1579 (1957).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 283–305.

ÜBER DEN EINFLUSS DER ANHARMONIZITÄT AUF DIE EIGENSCHAFTEN DER KRISTALLE*

WOLFGANG LUDWIG†

Institut für theoretische Physik der Universität Göttingen

(*Received* 13 *June* 1957)

Zusammenfassung—In der Gittertheorie der Kristalle beschränkt man sich im allgemeinen auf die harmonische Näherung, bei der in der Entwicklung der potentiellen Energie nach Verschiebungen nur quadratische Glieder berücksichtigt werden. In dieser Arbeit wird der Einfluss höherer Glieder (Anharmonizität) untersucht und an Hand eines speziellen Beispiels (kubisch-flächenzentriertes Gitter mit Zentralkräften nur zwischen nächsten Nachbarn) diskutiert.

Die freie Energie, aus der sich alle thermischen und kalorischen Grössen eines Kristalls ermitteln lassen, wird mit Hilfe eines der Diracschen Störungsrechnung analogen Verfahrens berechnet. Für die Temperaturabhängigkeit der isothermen und adiabatischen elastischen Konstanten lassen sich allgemeine Gesetze angeben. Die Cauchy-Relationen sind in keinem Fall erfüllt. Auch für die durch die Anharmonizität bewirkten Beiträge zur spezifischen Wärme lassen sich relativ einfache Ausdrücke angeben. Insbesondere wird gezeigt, dass der in der harmonischen Näherung abgeleitete Zusammenhang zwischen elastischen Daten und Debye-Temperatur auch am absoluten Nullpunkt nicht besteht. Es wird weiter ein Ausdruck für die Schallgeschwindigkeit bei Berücksichtigung der Anharmonizität angegeben. Die Beziehungen zwischen Schallgeschwindigkeit und elastischen Konstanten werden diskutiert. Es stellt sich heraus, dass für kubische Kristalle die in der harmonischen Theorie abgeleiteten Relationen nur für transversale Wellen richtig sind, während in anderen Fällen Abweichungen vorhanden sind. Das diskutierte einfache Beispiel würde zutreffen für feste Edelgase, für die aber nur wenig experimentelle Unterlagen vorliegen.

Abstract—In the lattice theory of crystals we generally assume the validity of the harmonic approximation, in which only quadratic terms are taken into account in the expansion of the potential energy in the displacements. In the present paper the influence of higher terms (anharmonicity) is investigated, and a particular example (face-centred cubic lattice with central forces between nearest neighbours only) is discussed.

The free energy, from which all the thermodynamic properties of a crystal may be evaluated, is computed with the aid of a method analogous to Dirac perturbation theory. It is possible to deduce general theorems for the temperature dependence of the isothermal and adiabatic elastic constants. In no case the Cauchy Relations are satisfied. It is also possible to derive relatively simple expressions for the contributions to the specific heat arising from anharmonicity. It is shown in particular, that the relation between elastic data and Debye temperature derived in the harmonic approximation is not valid even at absolute zero. An expression is also derived for the velocity of elastic waves taking into account anharmonicity. The relation between the wave velocity and the elastic constants is discussed. It appears from this that the relations deduced for cubic crystals are valid only for transverse waves. The simple example discussed would be valid for rare gases in the solid state, for which, however, only very scanty experimental data exist.

1. EINLEITUNG

EIN idealer Kristall wird aufgebaut aus regelmässig angeordneten Atomen oder Ionen, die Kräfte aufeinander ausüben. Aufgabe der Gittertheorie ist es, die elastischen, thermischen und kalorischen Daten eines solchen Kristalls zu berechnen und die Beziehungen zwischen ihnen aufzuzeigen.[1,2] Die exakte Behandlung dieses N-Teilchenproblems ist unmöglich. Wir gehen deshalb aus von der adiabatischen Näherung von BORN und OPPENHEIMER,[3] nach der sich für die Kerne eine potentielle Energie $\Phi(\ldots \boldsymbol{R}^m \ldots)$ angeben lässt. Ist diese

* Prof. M. BORN zum 75. Geburtstag am 11.12.57 gewidmet.

† D7. Jetzige Adresse: Lehrstuhl f. physikal. Grundlagen d. Reaktorwerkstoffe, Aachen, Techn. Hochschule, Templergraben 55.

als Funktion der Kernlagen $\boldsymbol{R}^m$ bekannt, sind wir in der Lage, alle thermischen und elastischen Eigenschaften eines Kristalls anzugeben. Die Gitterbausteine bewegen sich (bei nicht zu hohen Temperaturen und elastischen Beanspruchungen) nur wenig aus ihren mittleren Lagen* $\bar{\boldsymbol{R}}^m = A\boldsymbol{m}$; A ist die Matrix aus den drei Basisvektoren $\boldsymbol{a}^{(j)}$ des Gitters ($A_{ij} = a_i^{(j)}$) (Abb. 1), während $\boldsymbol{m}$ ein Vektor mit ganzzahligen Komponenten ist (im unendlichen Gitter!). Bezeichnen wir mit $\boldsymbol{q}^m$ die Verschiebung des $\boldsymbol{m}$-ten Atoms aus seiner mittleren Lage, so ist

$$\boldsymbol{R}^m = A\boldsymbol{m} + \boldsymbol{q}^m \tag{1.1}$$

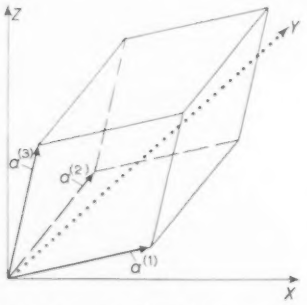


ABB. 1. Basisvektoren $\boldsymbol{a}^{(i)}$ und Elementarzelle eines primitiven Gitters.

Wir entwickeln nun die potentielle Energie nach diesen kleinen Verschiebungen um die Mittellagen, die im allgemeinen wegen der Anharmonizität nicht mit den durch das Minimum der potentiellen Energie definierten Lagen (Ruhelagen) übereinstimmen. Kenntnis der potentiellen Energie ist jetzt gleichbedeutend mit der Kenntnis der Entwicklungskoeffizienten, der sogenannten Kopplungsgrössen. Wir wollen sie im folgenden als bekannt voraussetzen und nur die Zusammenhänge zwischen ihnen und den thermischen und kalorischen Daten eines Kristalls betrachten.

Eine Reihe von Eigenschaften lässt sich schon im Rahmen der *harmonischen Näherung* erklären, bei der man die Entwicklung nach dem in den Verschiebungen quadratischen Glied abbricht :†

$$\Phi(\ldots\boldsymbol{R}^m\ldots) = \Phi(\ldots\bar{A}\boldsymbol{m}\ldots) + \sum_{mn} \tilde{\Phi}_{ij}^{mn} q_i^m q_j^n \tag{1.2}$$

Man hat es dann mit einem System von $3N$ Oszillatoren zu tun. Die Teilchen schwingen symmetrisch um ihre Gleichgewichtslage (Minimum der potentiellen Energie; deshalb tritt in (1.2) kein in den Verschiebungen lineares Glied auf); ihre Frequenz ist unabhängig von der Amplitude. Die Eigenfrequenzen $\tilde{\omega}_J$ wie auch alle makroskopischen Eigenschaften sind durch die atomistischen Kopplungsparameter $\tilde{\Phi}_{ij}^{mn}$ eindeutig festgelegt. Die freie Energie dieses Systems in Abhängigkeit von der Temperatur T und den Verzerrungen A–$\bar{A}$ (A beschreibt ein verzerrtes Gitter, während sich $\bar{A}$ auf die Ruhelage bezieht) lässt sich leicht angeben :

$$F = \Phi(\ldots A\boldsymbol{m}\ldots) + F_s;$$

$$F_s = kT \sum_J \ln\left[2\sinh\frac{\hbar\tilde{\omega}_J}{2kT}\right]. \tag{1.3}$$

Daraus kann man alle Grössen berechnen. Die $\tilde{\omega}_J$ hängen nicht von den Verzerrungen des Gitters ab (vgl.[1]). Die Gleichgewichtslagen und die elastischen Daten (thermische Zustandsgleichung) lassen sich aus $\Phi(A\boldsymbol{m})$ allein bestimmen, d.h. die elastische Theorie ist rein mechanisch und vollständig in Gleichung (1.2) enthalten. Es gibt daher keine thermische Ausdehnung und keine Temperaturabhängigkeit der elastischen Konstanten.

Die kalorischen Daten werden durch F_s allein bestimmt, da $\Phi(A\boldsymbol{m})$ temperaturunabhängig ist. Für die spezifische Wärme ($C_v = C_p$) ergibt sich bei hohen Temperaturen das Dulong-Petitsche Gesetz, wonach $C_v = 3NkT$ ist. Auch das Einmünden in diesen Grenzwert lässt sich angeben.[4,5] Bei tiefen Temperaturen benutzt man meist die Debyesche Näherung.[6] Danach ist

$$C_v = \frac{12\pi^4}{5} Nk(T/\Theta_D)^3. \tag{1.4}$$

Die auftretende charakteristische Temperatur (Debye-Temperatur) Θ_D lässt sich über die Schallgeschwindigkeiten aus den elastischen Daten des Kristalls berechnen. Beim Vergleich dieser Ergebnisse mit dem Experiment bedient man sich im allgemeinen des folgenden Verfahrens : die spezifische Wärme beschreibt man formal durch Gl. (1.4) und bestimmt den zu jeder Temperatur gehörigen Θ-Wert. Man erhält eine Funktion $\Theta(T)$. Wäre das T^3-Gesetz streng richtig, müsste Θ von der Temperatur unabhängig sein, was keineswegs der Fall ist. Bei höheren Temperaturen ist das nicht verwunderlich, da sich dort der Einfluss der Abweichungen des wahren Spektrums vom Debye-Spektrum bemerkbar macht.[4,5] Dagegen sollte im Limes $T \to 0$ $\Theta(T)$ gegen einen konstanten Wert streben, der mit dem aus den elastischen Daten berechneten übereinstimmt. Bei den meisten Kristallen zeigen sich selbst dann noch Diskrepanzen, die sich im Rahmen der harmonischen Näherung der Gittertheorie nicht erklären lassen.

* Wir beschränken uns auf primitive Gitter. Die Verallgemeinerung auf nicht primitive Gitter lässt sich ohne wesentliche Änderungen durchführen und soll in einer anderen Arbeit am Beispiel der Ionenkristalle diskutiert werden.

† Wir betrachten einen unendlichen Kristall, den wir in Periodizitätsvolumina von je N Atomen zerlegen. $\Phi(\ldots\boldsymbol{R}^m\ldots)$ ist die potentielle Energie eines dieser Volumina. Grössen, die sich auf das Minimum der potentiellen Energie als Ausgangslage beziehen, bezeichnen wir mit einer Tilde ($\sim$) ($\boldsymbol{R}^m = \bar{A}\boldsymbol{m} + \boldsymbol{q}^m$), nicht besonders gekennzeichnete beziehen sich auf allgemeine mittlere Lagen (wie in (1.1)).

VOL. 4 1958

Tatsächlich können wir gar nicht erwarten, dass die harmonische Theorie für tiefe Temperaturen richtig ist. Das Vorhandensein der Nullpunktsschwingungen bewirkt, dass selbst am absoluten Nullpunkt der Einfluss der Anharmonizität nicht verschwindet. Die mittleren Lagen $\boldsymbol{Am}$ stimmen keinesfalls mit dem Minimum der potentiellen Energie ($\tilde{A}\boldsymbol{m}$) überein (Abb. 2).

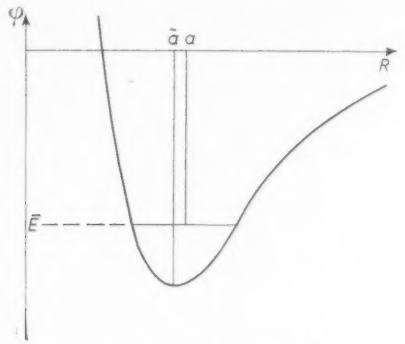


ABB. 2. Potentielle Energie zwischen zwei Atomen mit Ruhelagen $\tilde{a}$ und mittleren Lagen a.

Sobald die Anharmonizität eine grössere Rolle spielt,* enthält die elastische Theorie auch die Schwingungsenergie, d.h. sie ist nicht mehr reine Mechanik, sondern Thermodynamik. Alle statischen Daten (wie elastische Konstanten) erhält man direkt aus der freien Energie, während für dynamische Vorgänge (wie Schallwellen) die übliche Behandlung fraglich ist. Die Schallgeschwindigkeiten lassen sich sicher nicht mehr aus den statischen Daten allein bestimmen; damit ist auch der Zusammenhang zwischen Schallgeschwindigkeiten und elastischen Konstanten in Frage gestellt, und folglich auch der zwischen Debye-Temperatur und elastischen Daten. Was ist überhaupt die Schallgeschwindigkeit? Während elastische Daten und spezifische Wärmen nach wie vor aus der freien Energie bestimmt werden können, müssen wir uns bei Bestimmung der Schallgeschwindigkeit nach einer neuen Methode umsehen. Bei den Schallwellen handelt es sich um eine Überlagerung von mechanisch angeregten Schwingungen über die thermisch schon vorhandenen. Durch die Anharmonizität wird ein Energieaustausch zwischen Schwingungen verschiedener Frequenz bewirkt. Die mechanischen Schallwellen verlieren also auch Energie an die thermischen Schwingungen (und umgekehrt). Makroskopisch gesehen führt das zu einer Dämpfung der Schallwellen.

In der vorliegenden Arbeit sollen diese durch die Anharmonizität aufgeworfenen Fragen näher untersucht werden. Wir stellen alle Phänomene, die grundsätzlich erst durch Berücksichtigung der anharmonischen Glieder der potentiellen Energie erklärt werden können, noch einmal zusammen :

(1) Thermische Ausdehnung.

(2) Unterschied zwischen isothermen und adiabatischen elastischen Konstanten, ihre Temperaturabhängigkeit, sowie Abweichungen vom Hookeschen Gesetz bei grösseren Verformungen.

(3) Unterschied zwischen den spezifischen Wärmen C_v und C_p; ihre Temperaturabhängigkeit oberhalb der Debye-Temperatur, sowie die Bestimmung der Debye-Temperatur am absoluten Nullpunkt.

(4) Zusammenhang zwischen Schallgeschwindigkeit und elastischen Konstanten.

(5) Wärmeleitung in Isolatoren.

Der fünfte Punkt ist von LEIBFRIED und SCHLÖMANN abgehandelt worden.[7] Zu einem Teil der anderen Punkte haben Arbeiten von BORN und HOOTON[8,9] Beiträge geliefert. Wir wollen im folgenden einen anderen Weg beschreiten.

2. POTENTIELLE UND FREIE ENERGIE

Um die anharmonischen Effekte zu beschreiben, müssen wir also höhere Glieder in der Entwicklung der potentiellen Energie berücksichtigen. Wir schreiben

$$\Phi(\ldots \boldsymbol{R}^m \ldots) = \Phi_0 + \Phi_1 + \Phi_2 + \Phi_3 + \Phi_4 + \ldots \quad (2.1)$$

mit $\Phi_0 = \Phi(\ldots \boldsymbol{Am} \ldots)$; $\Phi_1 = \sum_{\substack{m \\ i}} \Phi_i^{\,m}\, q_i^{\,m}$;

$$\Phi_2 = \frac{1}{2!} \sum_{\substack{mn \\ ij}} \Phi_{ij}^{\,mn}\, q_i^{\,m}\, q_j^{\,n}; \qquad (2.1a)$$

$$\Phi_3 = \frac{1}{3!} \sum_{\substack{mno \\ ijk}} \Phi_{ijk}^{\,mno}\, q_i^{\,m}\, q_j^{\,n}\, q_k^{\,o};$$

$$\Phi_4 = \frac{1}{4!} \sum_{\substack{mnor \\ ijkl}} \Phi_{ijkl}^{\,mnor}\, q_i^{\,m}\, q_j^{\,n}\, q_k^{\,o}\, q_l^{\,r}$$

Die Entwicklung führen wir zweckmässig um die mittleren Lagen aus; diese spielen die Rolle von zunächst unbekannten Parametern, die nachträglich bestimmt werden aus der Forderung, dass die freie Energie F als Funktion von A minimal wird ($\partial F / \partial A_{ij} = 0$). Für die folgenden Rechnungen wollen wir annehmen, dass die Berücksichtigung der Glieder 4. Ordnung in (2.1) eine gute Näherung darstellt und höhere Glieder dagegen als klein vernachlässigt werden können. Man könnte zunächst meinen, in der Entwicklung (1.2) durch Hinzunahme des Gliedes Φ_3 auszukommen; es

* also vor allem bei kleinen Massen, wo die Nullpunktsunruhe gross ist.

wird sich aber herausstellen, dass die durch Φ_3 und Φ_4 bewirkten Effekte von gleicher Grösse sind. Entsprechend hätten wir bei besseren Näherungen Φ_5 *und* Φ_6 zu berücksichtigen. Die ungeraden Glieder bestimmen in der Hauptsache die statischen Eigenschaften (Gitterdaten, thermische Ausdehnung), die geraden Glieder die dynamischen (Eigenfrequenzen). Der Einfluss der Anharmonizität ist im allgemeinen um so grösser, je schwächer die Bindung eines Gitterbausteines an den Kristall ist und je kleiner seine Masse ist. Als charakteristische Grösse können wir etwa das Verhältnis von Bindungsenergie und Nullpunktsenergie ($\approx k\Theta$) ansehen, wobei Θ die Debye-Temperatur ist.

Das in den Verschiebungen lineare Glied in (2.1) verschwindet in primitiven Gittern, denn der Koeffizient Φ_i^m hat die Bedeutung der Kraft auf das Teilchen $\boldsymbol{m}$ in Richtung i, wenn alle Teilchen sich in ihren mittleren Lagen befinden. Aus Gründen der Gittersymmetrie darf dann aber keine Kraft auftreten, d.h. es ist $\Phi_i^m = 0$ für alle m und i.*

Ferner hängen die Grössen

$$\Phi_{i\,j}^{mn},\ \Phi_{i\,j\,k}^{mno},\ \ldots.$$

nur ab von $\boldsymbol{m}-\boldsymbol{n}$ bzw. $\boldsymbol{m}-\boldsymbol{o}$ und $\boldsymbol{n}-\boldsymbol{o}$; denn bei Verschiebung des gesamten Gitters um einen Basisvektor (Deckoperation des Gitters) dürfen sich die Kopplungsgrössen nicht ändern, also z.B.

$$\Phi_{i\quad j\quad k}^{m+h,n+h,o+h} = \Phi_{i\,j\,k}^{mno} = \Phi_{i\quad j\quad k}^{m-o,n-o\,0} \text{ für } \boldsymbol{h} = -\boldsymbol{o} \tag{2.2}$$

Der Anteil der Kraft auf das $\boldsymbol{m}$ -te Atom in Richtung i, der von einem bestimmten Entwicklungsglied herrührt, z.B. dem dritten, ist

$$K_i^m = -\tfrac{1}{2} \sum_{\substack{n\,o\\ j\,k}} \Phi_{i\,j\,k}^{mno}\, q_j^n q_k^o.$$

Da bei einer Verschiebung aller Atome um die gleiche Strecke aber keine Kraft auf das Atom $\boldsymbol{m}$ ausgeübt werden darf, so folgt

$$\sum_{n,o} \Phi_{i\,j\,k}^{mno} = 0, \text{ usw.} \tag{2.3}$$

Es sind nicht alle Kopplungsgrössen unabhängig voneinander, auch durch die Gittersymmetrien werden ihnen gewisse einschränkende Bedingungen auferlegt.[1]

In der harmonischen Theorie transformiert man die potentielle Energie (ebenso die kinetische) auf Normalkoordinaten a_s^k, und bekommt ein entkoppeltes System von Oszillatoren. Wir wollen auch hier die Transformation durchführen; natürlich bleibt über die Glieder dritter und vierter Ordnung in unserem Fall eine Kopplung bestehen. Die Beziehung zwischen Verschiebungen und Normalkoordinaten (stehende Wellen,[1,2]) ist

$$\boldsymbol{q}^m = \frac{1}{(NM)^{\frac{1}{2}}} \sum_{\substack{\boldsymbol{k}\\ s}} a_s^{\boldsymbol{k}} \boldsymbol{e}^s(\boldsymbol{k}) \exp(i\boldsymbol{k}A\boldsymbol{m}) \tag{2.4}$$

$\boldsymbol{e}^s(\boldsymbol{k})$ ist der Einheitsvektor der Polarisation s ($s = 1, 2, 3$) zur Wellenzahl $\boldsymbol{k}$, M die Masse der Gitterbausteine. Wir erhalten schliesslich für die Hamiltonfunktion

$$\mathscr{H} = \mathscr{H}_0 + \Phi_3 + \Phi_4 \tag{2.5}$$

mit

$$\mathscr{H}_0 = \Phi_0(\ldots A\boldsymbol{h}\ldots) + \\ + \tfrac{1}{2} \sum_{\substack{\boldsymbol{k}\\ s}} \{\dot{a}_s^{*\boldsymbol{k}} \dot{a}_s^{\boldsymbol{k}} + \omega_s^2(\boldsymbol{k}) a_s^{*\boldsymbol{k}} a_s^{\boldsymbol{k}}\}$$

$$\Phi_3 = \frac{1}{3!} \cdot \frac{1}{N^{\frac{1}{2}}} \sum_{\substack{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}''\\ s\,s'\,s''}} \Phi_{s\,s'\,s''\,s'''}^{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}''\boldsymbol{k}'''} a_s^{\boldsymbol{k}} a_{s'}^{\boldsymbol{k}'} a_{s''}^{\boldsymbol{k}''} a_{s'''}^{\boldsymbol{k}'''} \tag{2.6}$$

$$\Phi_4 = \frac{1}{4!} \cdot \frac{1}{N} \sum_{\substack{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}''\boldsymbol{k}'''\\ s\,s'\,s''\,s'''}} \Phi_{s\,s'\,s''\,s'''}^{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}''\boldsymbol{k}'''} a_s^{\boldsymbol{k}} a_{s'}^{\boldsymbol{k}'} a_{s''}^{\boldsymbol{k}''} a_{s'''}^{\boldsymbol{k}'''}.$$

Die $\omega_s(\boldsymbol{k})$ bestimmen sich aus der Säkulargleichung

$$\omega_s^2(\boldsymbol{k}) e_i^s(\boldsymbol{k}) = \sum_j t_{ij}(\boldsymbol{k}) e_j^s(\boldsymbol{k}) \text{ mit} \tag{2.7a}$$

$$t_{ij}(\boldsymbol{k}) = \frac{1}{M} \sum_{\boldsymbol{h}} \Phi_{ij}^{\boldsymbol{h}} e^{-i\boldsymbol{k}A\boldsymbol{h}}$$

* Auch in nicht-primitiven Gittern kann dieses Glied weitgehend unterdrückt werden, so dass unsere späteren Resultate fast unmittelbar übertragen werden können. Wir wollen darauf hier nicht näher eingehen.

VOL. 4 1958

Ferner ist

$$\Phi^{\boldsymbol{kk'k''}}_{ss's''} = \frac{1}{N\cdot M^{3/2}} \sum_{\substack{mno\\ijk}} \Phi^{mno}_{ijk} \times$$

$$\times e^{s}_{i}(\boldsymbol{k}) e^{s'}_{j}(\boldsymbol{k}') e^{s''}_{k}(\boldsymbol{k}'') \exp[i(\boldsymbol{kAm}+\boldsymbol{k}'\boldsymbol{An}+\boldsymbol{k}''\boldsymbol{Ao})] \tag{2.7b}$$

$$\Phi^{\boldsymbol{kk'k''k'''}}_{ss's''s'''} = \frac{1}{N\cdot M^{2}} \sum_{\substack{mnor\\ijkl}} \Phi^{mnor}_{ijkl} \times$$

$$\times e^{s}_{i}(\boldsymbol{k}) e^{s'}_{j}(\boldsymbol{k}') e^{s''}_{k}(\boldsymbol{k}'') e^{s'''}_{l}(\boldsymbol{k}''') \times$$

$$\times \exp[i(\boldsymbol{kAm}+\boldsymbol{k}'\boldsymbol{An}+\boldsymbol{k}''\boldsymbol{Ao}+\boldsymbol{k}'''\boldsymbol{Ar})].$$

Man überzeugt sich leicht davon, dass die $\Phi^{\boldsymbol{kk'}\ldots}_{ss'\ldots}$ gegen Vertauschungen der Indexpaare $\begin{matrix}\boldsymbol{k}\\ s\end{matrix}, \begin{matrix}\boldsymbol{k'}\\ s'\end{matrix}, \ldots$ invariant sind.

Wir wollen hier noch eine kurze Bemerkung anschliessen über die Argumente der in (2.7) auftretenden Exponentialfunktionen, z.B. $\boldsymbol{kAm}$. Es könnte auf den ersten Blick so scheinen, als ob sie von dem Parameter A abhingen. Nun gilt aber für die Wellenzahlen $\boldsymbol{k}$ auf Grund der angenommenen Periodizität des Gitters*

$$\boldsymbol{k} = \frac{2\pi}{\sqrt[3]{N}} \boldsymbol{Bn} \quad \text{mit} \quad B'A = 1$$

also ist

$$\boldsymbol{kAm} = \frac{2\pi}{\sqrt[3]{N}} (\boldsymbol{Bn}, \boldsymbol{Am}) = \frac{2\pi}{\sqrt[3]{N}} (\boldsymbol{n}, \boldsymbol{m})$$

eine reine Zahl, unabhängig von dem Parameter A. Dieser kommt in (2.7) demnach nur in den Kopplungsgrössen $\Phi^{\boldsymbol{mn}\ldots}_{ij\ldots}$ vor. Vielleicht sollte man deshalb die $\Phi^{\boldsymbol{kk'}\ldots}_{ss'\ldots}$ statt mit $\boldsymbol{k}$ besser mit einer ganzen Zahl $(\boldsymbol{n})$ indizieren, doch hat $\boldsymbol{k}$ den Vorzug einer anschaulicheren Bedeutung.

Zur Berechnung der freien Energie bedienen wir uns einer konsequenten Entwicklung der Zustandssumme Z, wie sie von LEIBFRIED [1] und von NAKAJIMA[10] angegeben worden ist. Hat der Hamiltonoperator die Gestalt $\mathscr{H} = \mathscr{H}_0 + gW$, so wird die freie Energie†

$$-F/kT = \ln Z(0) + \frac{Z'(0)}{Z(0)} + \frac{1}{2}\left\{\frac{Z''(0)}{Z(0)} - \left(\frac{Z'(0)}{Z(0)}\right)^2\right\} + \ldots \tag{2.8}$$

mit

$$Z(0) = \sum_{\alpha} \exp(-\beta E_\alpha);$$

$$Z'(0) = -\beta \sum_{\alpha} W_{\alpha\alpha} \exp(-\beta E_\alpha);$$

$$Z''(0) = 2 \sum_{\alpha,\gamma} W_{\alpha\gamma} W_{\gamma\alpha} \times \frac{\exp[\beta(E_\alpha - E_\gamma)] - 1 - \beta(E_\alpha - E_\gamma)}{(E_\alpha - E_\gamma)^2} \exp(-\beta E_\alpha);$$

$$\beta = \frac{1}{kT} \tag{2.8a}$$

Als Funktionensystem ist das des ungestörten Hamiltonoperators $\mathscr{H}_0$ benutzt worden. E_α ist der Eigenwert von $\mathscr{H}_0$ im Zustand α, $W_{\alpha\gamma}$ das Matrixelement der Störung, gebildet mit den Zuständen α und γ. Diese Entwicklung entspricht der Diracschen Störungsrechnung 2. Näherung. Der Hamiltonoperator unseres ungestörten Problems $\mathscr{H}_0$, ist der eines Systems von unabhängigen Oszillatoren nach (2.6). Die Störung ist $W = \Phi_3 + \Phi_4$.

Man sieht sofort folgendes: Die freie Energie nullter Näherung ist dieselbe wie in der harmonischen Näherung, abgesehen davon, dass die Frequenzen hier von den noch zu bestimmenden Parametern A abhängen. In erster Näherung gehen nur die Diagonalelemente $W_{\alpha\alpha}$ der Störung ein. Nun enthält aber der von Φ_3 herrührende Anteil keine Diagonalelemente, da die Koordinaten in ungerader Potenz auftreten; in erster

* Das Periodizitätsvolumen sei so gewählt, dass es in Richtung aller drei Basisvektoren gleich viel Atome enthalte, also $\sqrt[3]{N}$.

† Die Grösse $F/kT = -\ln Z(g)$ wird nach Potenzen von g entwickelt ($Z'(0) = \partial(g)/Z\partial g|_{g=0}$; ...). Der Parameter g dient nur zur Indizierung der Potenzen von W. g wird daher nach der Ausführung der Entwicklung gleich Eins gesetzt.

Näherung gibt nur Φ_4 einen Beitrag zur freien Energie, den wir mit F_4 bezeichnen wollen:

$$F_4 = -\frac{1}{\beta} \cdot \frac{Z'(0)}{Z(0)}. \tag{2.9a}$$

Erst in zweiter Näherung liefert Φ_3 einen Beitrag (F_3), der natürlich die Kopplungsgrössen 3. Ordnung quadratisch enthält. Ebenso gibt Φ_4 in zweiter Näherung einen Beitrag. Diesen wollen wir vernachlässigen; denn er liefert Terme, die von höherer Ordnung klein sind.* Dagegen sind die linearen Beiträge von Φ_4 und die quadratischen von Φ_3 etwa von gleicher Grösse. Höhere Glieder als diese sollen nicht berücksichtigt werden. Wir haben demnach

$$F_3 = -\frac{1}{2\beta} \cdot \frac{Z''(0)}{Z(0)}. \tag{2.9b}$$

Die etwas langwierige, aber streng mögliche Berechnung von F_3 und F_4 wollen wir hier nicht durchführen (vgl. Anhang 1). Die freie Energie wird insgesamt

$$F = \Phi_0 + F_s + F_3 + F_4 \tag{2.10}$$

mit

$$F_s = kT \cdot \sum_{\boldsymbol{k},s} \ln[2 \sinh \tfrac{1}{2}\beta\hbar\omega_s(\boldsymbol{k})] \tag{2.10a}$$

$$F_3 = -\frac{\hbar^2}{16N} \sum_{\substack{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}'' \\ s\,s'\,s''}} \frac{|\Phi^{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}''}_{s\,s'\,s''}|^2}{\omega\omega'\omega''} \times$$

$$\times \left\{ \frac{\bar{n} + \bar{n} \cdot \bar{n}'}{\omega + \omega' + \omega''} + \frac{\bar{n}'' + 2\bar{n} \cdot \bar{n}'' - \bar{n} \cdot \bar{n}'}{\omega + \omega' - \omega''} \right\} \tag{2.10b}$$

$$F_4 = \frac{\hbar}{8N} \sum_{\substack{\boldsymbol{k}\boldsymbol{k}' \\ s\,s'}} \frac{\Phi^{\boldsymbol{k}-\boldsymbol{k}\boldsymbol{k}'-\boldsymbol{k}'}_{s\;s\;s'\;s'}}{\omega\omega'} (\bar{n} + \tfrac{1}{2})(\bar{n}' + \tfrac{1}{2}).$$

Zur Abkürzung ist gesetzt worden : $\omega = \omega_s(\boldsymbol{k})$; $\omega' = \omega_{s'}(\boldsymbol{k}')$; ... . Die Grösse $\bar{n} = 1/(e^{\beta\hbar\omega} - 1)$ ist die mittlere Besetzungszahl eines Planckschen Oszillators.

* Sie wird von gleicher Grössenordnung wie die Beiträge, die von den schon oben vernachlässigten Gliedern Φ_5 und Φ_6 herrühren.

Alle vier angegebenen Terme der freien Energie hängen zunächst noch von den Parametern A ab, über die wir folgendes aussagen können. Die Differenz zwischen den Parametern A und den durch die Ruhelagen definierten $\tilde{A}$ muss proportional sein zu den Kopplungsgrössen 3. Ordnung (oder zu Potenzen davon oder zu Kopplungsgrössen noch höherer Ordnung), denn in der harmonischen Näherung fallen A und $\tilde{A}$ zusammen. Da nun F_3 und F_4 schon proportional sind zu Φ^{mno}_{ijk} bzw. Φ^{mnor}_{ijkl}, und wir noch höhere Effekte vernachlässigen, so können wir in F_3 und F_4 überall A durch $\tilde{A}$ ersetzen. F_3 und F_4 hängen nicht mehr von den Verzerrungen aus den Ruhelagen ab, nur noch von der Temperatur. Sie geben demnach keinen Beitrag zur thermischen Ausdehnung und den elastischen Konstanten (Ableitungen der freien Energie nach den Verzerrungen), wohl aber zur spezifischen Wärme. In F_3 und F_4 können wir in den ω *und* in Φ_3 A durch $\tilde{A}$ ersetzen, da $A-\tilde{A}$ quadratisch eingeht. Φ_4 dagegen enthält schon auf Grund unseres Modells nur noch $\tilde{A}$.

3. DIE THERMISCHE ZUSTANDSGLEICHUNG

Bevor wir weitere Aussagen über die Eigenschaften der Kristalle machen können, müssen wir die mittleren Lagen $\boldsymbol{Am}$ bestimmen aus dem Minimum der freien Energie :

$$\left(\frac{\partial F}{\partial A_{ij}}\right)_T = 0. \tag{3.1}$$

Anstelle der A_{ij} rechnet man bequemer mit den Verzerrungen V_{ij} aus den Ruhelagen $\tilde{A}_{ij}$. Sie sind definiert durch (Abb. 3):

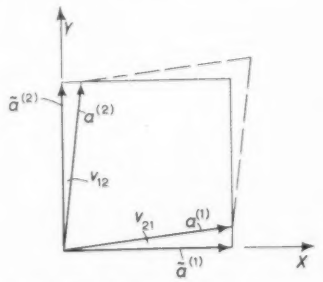


ABB. 3. Zur Definition einer Verzerrung V_{ij}.

$$\boldsymbol{A} = (1+V)\,\tilde{A} \text{ bzw. } A_{ij} = \sum_k (\delta_{ij} + V_{ik})\tilde{A}_{kj}. \tag{3.2}$$

VOL. 4 1958

V ist der Tensor der endlichen Verzerrungen. Speziell soll die durch die Temperatur bewirkte Verzerrung (thermische Ausdehnung) mit V_{ij}^{th} bezeichnet werden. Die Forderung (3.1) ist dann gleichbedeutend mit

$$(\partial F/\partial V_{ij})_{T;\,V_{ij}=V_{ij}^{\text{th}}}=0. \tag{3.3}$$

Diese Gleichung ist gleichbedeutend mit dem Verschwinden der äusseren Spannungen in dem durch T und V_{ij}^{th} gegebenen Zustand. Das folgt aus der Thermodynamik der Kristalle.[1] Der Einfluss des Dampfdrucks ist dabei vernachlässigt.

Aus (2.10) und (2.10a) erhalten wir (vgl. Anhang 2)*

$$V_{jl}^{\text{th}}=-\frac{1}{2NV_z}\sum_{ik}C^{-1}{}_{jl,ik}\left(\frac{\partial\ln\overline{\omega^2}}{\partial V_{ik}}\right)_{V=V^{\text{th}}}\cdot U_h \tag{3.4a}$$

bzw. den Tensor des thermischen Ausdehnungskoeffizienten

$$\alpha_{jl}=-\frac{1}{2NV_z}\sum_{ik}C^{-1}{}_{jl,ik}\left(\frac{\partial\ln\overline{\omega^2}}{\partial V_{ik}}\right)_{V=V^{\text{th}}}\cdot C_v^h \tag{3.4b}$$

$\overline{\omega^2}$ ist der Mittelwert von $\omega_s^2(\boldsymbol{k})$ über alle vorkommenden Wellenzahlen $\boldsymbol{k}$ und Polarisationsrichtungen s. V_z ist das Volumen einer Elementarzelle, die $C_{ik,jl}$ sind die elastischen Konstanten, U_h und C_v^h innere Energie und spezifische Wärme, alles in der harmonischen Näherung, d.h. diese Grössen sind alle auf die Ruhelagen $\tilde{A}$ bezogen.

Im kubischen Fall ist die Ausdehnung isotrop; die Gln. (3.4a) und (3.4b) lassen dann einige Vereinfachungen zu.
Es wird $V_{jl}^{\text{th}}=V_0^{\text{th}}\cdot\delta_{jl}$ mit

$$V_0^{\text{th}}=\frac{\gamma\kappa}{3NV_z}\cdot U_h \text{ und } \alpha=\frac{\gamma\kappa}{3NV_z}\cdot C_v^h. \tag{3.5}$$

Hier ist κ die Kompressibilität der harmonischen Näherung, γ ist die durch $\partial\ln\overline{\omega^2}/\partial V_z=-2\gamma/V_z$ definierte Grüneisensche Konstante.[11] Sie ist in dieser Näherung temperaturunabhängig.

Wir sind nun in der Lage, die elastischen Konstanten† aus den thermodynamischen Beziehungen

$$C_{ik,j}^{\text{is}}=\frac{1}{NV_z}\cdot\left(\frac{\partial^2F}{\partial V_{ik}\partial V_{jl}}\right)_{T;\,V_{ik}=V_{ik}^{\text{th}}}; \tag{3.6}$$

$$C_{ik,jl}^{\text{ad}}-C_{ik,jl}^{\text{is}}=\frac{T}{NV_zC_v^h}\cdot\left(\frac{\partial^2F}{\partial T\partial V_{ik}}\cdot\frac{\partial^2F}{\partial T\partial V_{jl}}\right)_{V_{ik}=V_{ik}^{\text{th}}}$$

zu berechnen. Es ergibt sich

$$C_{ik,jl}^{\text{ad}}=\frac{1}{N}\cdot\left(\frac{\partial^2\Phi(Am)}{\partial V_{ik}\partial V_{jl}}\right)_{T;\,V_{ik}=V_{ik}^{\text{th}}}+$$
$$+\frac{1}{NV_z}\left(\tfrac{1}{2}\frac{\partial^2\ln\overline{\omega^2}}{\partial V_{ik}\partial V_{jl}}+\tfrac{1}{4}\frac{\partial\ln\overline{\omega^2}}{\partial V_{ik}}\frac{\partial\ln\overline{\omega^2}}{\partial V_{jl}}\right)\cdot U_h \tag{3.7a}$$

und

$$C_{ik,jl}^{\text{ad}}-C_{ik,jl}^{\text{is}}=\frac{1}{4NV_z}\cdot\frac{\partial\ln\overline{\omega^2}}{\partial V_{ik}}\frac{\partial\ln\overline{\omega^2}}{\partial V_{jl}}\cdot TC_v^h. \tag{3.7b}$$

Der erste Anteil in (3.7a) enthält einmal die elastischen Konstanten der harmonischen Näherung, zum anderen aber auch einen Anteil, der proportional zu U_h ist (vgl. Anhang 2, A.6). Wir wollen die explizite Form hier nicht angeben, da wir später die elastischen Konstanten für einen speziellen Fall ableiten. Wesentlich ist, dass die adiabatischen Konstanten in ihrer Temperaturabhängigkeit proportional zur inneren Energie U_h sind,

* Die zunächst auftretenden Summen der Form

$$\frac{1}{3N}\sum_{k,s}f(\omega_s^2(\boldsymbol{k}))$$

sind hier durch $f(\overline{\omega_s^2})$ ersetzt worden. $C^{-1}{}_{jl,ik}$ ist das Reziproke des (vierstufigen) Tensors der elastischen Konstanten $C_{ik,jl}$.

† Die elastischen Konstanten unserer Näherung sind immer mit einem Index (is bzw. ad) versehen, während die Konstanten der harmonischen Näherung keine nähere Bezeichnung tragen.

Anm. b.d. Korrektur: E. A. STERN gibt in *Bull. Amer. Phys. Soc.* II, **2**, 214 (57) eine andere Methode zur direkten Berechnung der adiabatischen Konstanten an, die natürlich dasselbe Ergebnis liefert.

während sich die isothermen von ihnen durch einen Term unterscheiden, der proportional zu TC_v^h ist; der Unterschied verschwindet für $T=0$. In beiden Fällen erreichen die elastischen Konstanten für $T=0$ nicht den Wert der harmonischen Näherung. Diesen Wert erhält man vielmehr sowohl aus den adiabatischen wie den isothermen elastischen Konstanten, wenn man den bei hohen Temperaturen gemessenen Verlauf linear auf $T=0$ extrapoliert, da dort sowohl U_h als auch TC_v^h proportional zu T sind.*

Die Gleichungen (3.7a) und (3.7b) lassen bei Berücksichtigung der Gittersymmetrie natürlich noch viele Vereinfachungen zu. Bei kubischer Symmetrie gilt z.B. für $C_{12,12}^{\text{is}} = c_{44}^{\text{is}}$ (Voigtsche Indizierung)†:

$$c_{44}^{\text{is}} = c_{44}^{\text{ad}} = \frac{1}{N}\frac{\partial^2\Phi(Am)}{\partial V_{12}^2} + \frac{1}{2NV_z}\cdot\frac{\partial^2 \ln\overline{\omega^2}}{\partial V_{12}^2}\cdot U_h. \tag{3.8a}$$

Für die Kompressibilität erhalten wir nach etwas längerer Rechnung :

$$\frac{1}{\kappa^{\text{ad}}} = \frac{1}{\kappa} + \frac{\kappa\gamma}{27N^2V_z}\sum_{\substack{mno\\ijk}}\tilde{\Phi}_{ijk}^{mno}(\tilde{A}m)_i(\tilde{A}n)_j(\tilde{A}o)_k U_h + \frac{1}{NV_z}\left\{2V_z\frac{\partial\gamma}{\partial V_z}+\gamma^2\right\}\cdot U_h$$

und

$$\frac{1}{\kappa^{\text{ad}}} - \frac{1}{\kappa^{\text{is}}} = \frac{\gamma^2}{NV_z}\cdot TC_v^h.$$

Wir wollen die hier angegebenen Zusammenhänge etwas erläutern für den einfachsten Fall, der zu einem stabilen Kristall führt : ein kubisch-flächenzentriertes Gitter mit Zentralkräften nur zwischen nächsten Nachburn. Die potentielle Energie zwischen zwei Atomen ist :

$$\phi(R) = \phi(\tilde{l}) + \tfrac{1}{2}f(R-\tilde{l})^2 + \frac{1}{6}g(R-\tilde{l})^3 + \frac{1}{24}h(R-\tilde{l})^4 + \ldots \tag{3.9}$$

$R=\tilde{l}$ ist das Minimum der potentiellen Energie. Das Modell enthält drei wesentliche Parameter : die Federkonstante f und die Anharmonizitätskonstanten g und h. Die zum Teil etwas längeren Rechnungen wollen wir übergehen und nur das Ergebnis angeben (a : kubische Gitterkonstante; $V_z = a^3/4$) :

$$V_0^{\text{th}} = -\frac{\sqrt{2}}{12f^2aN}(g+2\sqrt{2}f/a)U_h;$$
$$\alpha = -\frac{\sqrt{2}}{12f^2aN}(g+2\sqrt{2}f/a)C_v^h; \tag{3.10}$$
$$\gamma = -\frac{ga}{6\sqrt{2}f}\left(1+\frac{2\sqrt{2}f}{ga}\right)$$

Für die elastischen Konstanten ergibt sich :

$$c_{11}^{\text{is}} = \frac{2f}{a}\left\{1 - \frac{1}{12f^2N}(\tfrac{4}{3}g^2/f - h + \sqrt{2}/3\cdot g/a + 20/3\cdot f/a^2)U_h - \frac{1}{36f^3N}(g+2\sqrt{2}f/a^2)^2 TC_v^h\right\}$$

$$c_{44}^{\text{is}} = \frac{f}{a}\left\{1 - \frac{1}{12f^2N}(g^2/f - h - \sqrt{2}g/a + 4f/a^2)U_h\right\} \tag{3.11}$$

$$c_{12}^{\text{is}} = c_{44}^{\text{is}} - \frac{1}{18f^2aN}(g+2\sqrt{2}f/a)^2(U_h + TC_v^h)$$

$$= c_{11}^{\text{is}} - c_{44}^{\text{is}}$$

$$\frac{1}{\kappa^{\text{is}}} = \frac{c_{11}^{\text{is}} + 2c_{12}^{\text{is}}}{3}.$$

Die adiabatischen Konstanten erhalten wir hieraus gemäss der allgemein angegebenen Vorschrift durch Streichung des Gliedes mit TC_v^h. Das hier $c_{44}^{\text{is}} = c_{11}^{\text{is}} - c_{12}^{\text{is}}$ ist, liegt natürlich an unserem speziellen Modell. Der Kristall ist elastisch anisotrop, das Verhältnis der beiden Schubmoduln (c_{44}^{is} und $\frac{1}{2}(c_{11}^{\text{is}} - c_{12}^{\text{is}})$) ist zwei. Wir erkennen ferner sofort, dass die Cauchy-Relation

* Für hohe Temperaturen ist die Temperaturabhängigkeit der elastischen Konstanten für einen speziellen Fall schon von Gow und Born[12] angegeben worden (Vgl. auch Literatur (1)).

† Die Gleichheit von c_{44}^{is} und c_{44}^{ad} folgt natürlich schon aus der allgemeinen Thermodynamik der Kristalle.

VOL. 4 1958

$c_{44}^{is} = c_{12}^{is}$ nicht erfüllt ist, obwohl wir reine Zentralkräfte angenommen haben. Auch am absoluten Nullpunkt sind wegen der immer vorhandenen Nullpunktsenergie Abweichungen von der Cauchy-Relation vorhanden. Erst wenn wir die bei hohen Temperaturen gemessenen Konstanten linear auf $T = 0$ extrapolieren (harmonischer Grenzfall), ist $c_{44} = c_{12}$. Die Abweichungen werden natürlich mit wachsender Anharmonizität grösser. Für spezielle Werte von g^2/f, h, g/a und f/a^2 kann es sein, dass die elastischen Konstanten nahezu temperaturunabhängig sind; dann müssen jedoch, falls nur Zentralkräfte wirksam sind, die Cauchy-Relationen erfüllt sein. Es sei noch auf zwei Tatsachen hingewiesen, die sich aus den allgemeinen Formeln leicht herauspräparieren lassen : einmal die relative Temperaturunempfindlichkeit von c_{44} (c_{12} und c_{11} enthalten zwei temperaturabhängige Anteile mehr, vgl. (3.7) und (3.8a)), zum anderen sollte bei Zentralkräften für alle Temperaturen c_{44} stets grösser als c_{12} sein. Die Cauchy-Relationen sind natürlich in keinem Fall erfüllt.

Das vorstehend diskutierte Modell sollte am ehesten zutreffend sein für die Edelgase im festen Zustand. Die potentielle Energie zwischen zwei Atomen lässt sich schreiben als*

$$\phi(R) = -4\epsilon\left\{\left(\frac{\sigma}{R}\right)^6 - \left(\frac{\sigma}{R}\right)^{12}\right\}$$

Es wird dann mit

$$f = 57{,}5\epsilon/\sigma^2;\ g = -1072\epsilon/\sigma^3,\ h = 16804\epsilon/\sigma^4$$

$$c_{11}^{is} = \frac{2f}{a}\left\{1 - 0{,}25U_h/N\epsilon - 0{,}14TC_v^{\hbar}/N\epsilon\right\}$$

$$c_{12}^{is} = \frac{f}{a}\left\{1 - 0{,}39U_h/N\epsilon - 0{,}28TC_v^{\hbar}/N\epsilon\right\}$$

$$c_{44}^{is} = \frac{f}{a}\left\{1 - 0{,}11U_h/N\epsilon\right\} = c_{11}^{is} - c_{12}^{is}$$

$$\frac{1}{\kappa^{is}} = \frac{4f}{3a}\left\{1 - 0{,}32U_h/N\epsilon - 0{,}21TC_v^{\hbar}/N\epsilon\right\}.$$

* Als Abstossung wählen wir der Einfachheit halber $1/R^{12}$. Andere Potenzen oder ein exponentialles Abstossungsgesetz liefern praktisch dasselbe.

Der Verlauf ist mit dem experimentell ermittelten Wert von ϵ für Krypton in Abb. 4 dargestellt. In dem angenommenen Modell nächster Nachbar-Wechselwirkung ist ϵ die Bindungsenergie eines Atoms, $a = 2^{\frac{1}{2}}\sigma$ die Gitterkonstante. In der Grössenordnung stimmt der numerische Wert

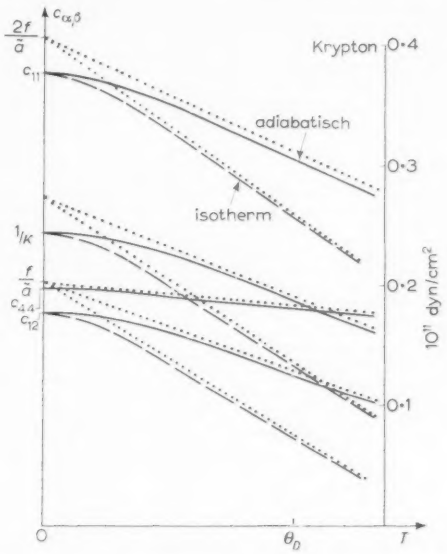


ABB. 4. Die Temperaturabhängigkeit der elastischen Konstanten für kubisch-flächenzentriertes Gitter mit Zentralkräften nur zwischen nächsten Nachbarn. Die Zahlen auf der rechten Seite entsprechen den Werten für Krypton in unserem speziellen Modell (vgl. Text).
ausgezogen : adiabatische Konstanten
gestrichelt : isotherme Konstanten
punktiert : lineare Extrapolation

für die Kompressibilität mit dem gemessenen überein; der Temperaturverlauf ist leider nicht gemessen. Eine vollkommene numerische Übereinstimmung dürfen wir nicht erwarten, da wir nur nächste Nachbarn berücksichtigt haben. Der Einfluss weiterer Nachbarn beträgt immerhin 15 %, in einigen Effekten, wo er quadratisch eingeht, 30 %. Diesen Einfluss kann man selbstverständlich leicht noch berücksichtigen, was an dieser Stelle unterbleiben soll.

4. DIE KALORISCHE ZUSTANDSGLEICHUNG

Für alle kalorischen Daten sind die Terme F_3 und F_4 der freien Energie wesentlich. In der

Hauptsache interessiert uns nun die spezifische Wärme (C_v, C_p). Das Glied $\Phi(...Ah...)$ liefert keinen Beitrag zu C_v, denn es hängt nur über die mittleren Lagen von der Temperatur ab, die bei der Bildung von C_v aber gerade konstant gehalten werden. Nur die explizite Temperaturabhängigkeit darf bei der Berechnung von C_v berücksichtigt werden. Die spezifische Wärme bei konstantem Druck, C_p, bestimmen wir aus den allgemeinen thermodynamischen Relationen :

$$C_p - C_v = -T \sum_{ik} \frac{\partial^2 F}{\partial T \partial V_{ik}} \cdot \alpha_{ik}$$

$$= TNV_z \sum_{ik,jl} \alpha_{ik} C_{ik,jl} \alpha_{jl} \qquad (4.1)$$

bzw. speziell für kubische Kristalle

$$C_p - C_v = TNV_z \cdot \frac{9\alpha^2}{\kappa}. \qquad (4.1a)$$

Wegen (3.4b) ist die Differenz $C_p - C_v$ proportional zu TC_v^{h2}, also bei tiefen Temperaturen $\sim T^7$, bei hohen $\sim T$.

Schon in der harmonischen Theorie ist es schwierig, den Ausdruck (2.10a) für den ganzen Temperaturbereich geschlossen zu diskutieren. Die Untersuchung der anharmonischen Terme (2.10b) ist naturgemäss noch umständlicher. Wir betrachten die Fälle hoher ($T > \Theta$) und tiefer Temperaturen gesondert.

(a) *Hohe Temperaturen*

Der harmonische Anteil (2.10a) liefert in bekannter Weise :

$$F_s = kT \sum_{\boldsymbol{k},s} \ln \frac{\hbar\omega_s(\boldsymbol{k})}{kT} +$$

$$+ \frac{1}{8} \frac{\hbar^2 \overline{\omega^2}}{(kT)^2} - \frac{1}{960} \frac{\hbar^4 \overline{\omega^4}}{(kT)^4} \qquad (4.2)$$

F_3 und F_4 lassen sich genau so entwickeln und geben nach einfacher Rechnung* :

$$F_3 = -\frac{(kT)^2}{12N} \sum_{\substack{\boldsymbol{kk'k''} \\ ss's''}} \frac{|\Phi^{\boldsymbol{kk'k''}}_{ss's''}|^2}{\omega^2 \omega'^2 \omega''^2} \times$$

$$\times \left\{ 1 + \frac{\hbar^4 \omega^2 \omega'^2}{240(kT)^4} \right\} \qquad (4.3a)$$

$$F_4 = \frac{(kT)^2}{8N} \sum_{\substack{\boldsymbol{kk'} \\ ss'}} \frac{\Phi^{\boldsymbol{k}-\boldsymbol{kk'}-\boldsymbol{k'}}_{s\ s\ s'\ s'}}{\omega^2 \omega'^2} \times$$

$$\times \left\{ \frac{\hbar^2 \omega^2}{6(kT)^2} - \frac{\hbar^4 \omega^2 [2\omega^2 - 5\omega'^2]}{720(kT)^4} \right\}. \qquad (4.3b)$$

Um diese Summen auswerten zu können, machen wir folgende Näherung : Wir ziehen alle Potenzen von $\omega_s^2(\boldsymbol{k})$ als Mittelwerte vor das Summenzeichen, in (4.3a) erhalten wir bei dem ersten Term dann z.B. $(1/\overline{\omega^2})^3$. Man könnte zunächst meinen, es als $(\overline{1/\omega^2})^3$ vorziehen zu müssen. Man muss aber beachten, dass auch die $\Phi^{\boldsymbol{kk'k''}}_{ss's''}$ noch von $\boldsymbol{k}$ abhängen, und zwar in ähnlicher Weise wie die $\omega_s^2(\boldsymbol{k})$, so dass sich in Wirklichkeit die $\boldsymbol{k}$-Abhängigkeit von Zähler und Nenner teilweise gegenseitig kompensieren. Das wird besser berücksichtigt, wenn man $(1/\overline{\omega^2})^3$ herauszieht. An dem einfachen Fall der linearen Kette lassen sich alle Grössen exakt berechnen; die Resultate werden wesentlich besser durch $(1/\overline{\omega^2})^3$ wiedergegeben. Die restlichen Summen $\sum_{\substack{\boldsymbol{kk'k''} \\ ss's''}} |\Phi^{\boldsymbol{kk'k''}}_{ss's''}|^2$ und $\sum_{\substack{\boldsymbol{kk'} \\ ss'}} \Phi^{\boldsymbol{k}-\boldsymbol{kk'}-\boldsymbol{k'}}_{s\ s\ s'\ s'}$ lassen sich sofort ausführen und wir erhalten schliesslich für die spezifische Wärme :

$$C_v/3Nk = 1 + \frac{kT}{6(M\overline{\omega^2})^2} \times$$

$$\times \left\{ \frac{1}{3M\overline{\omega^2}} \sum_{\substack{mn \\ ijk}} \Phi^{mn0^2}_{ijk} - \frac{1}{2} \sum_{\substack{m \\ ij}} \Phi^{mm00}_{iijj} \right\} -$$

$$- \frac{1}{12} \frac{\hbar^2 \overline{\omega^2}}{(kT)^2} + \frac{1}{240} \frac{\hbar^4 \overline{\omega^4}}{(kT)^4} \times \qquad (4.5)$$

$$\times \left\{ 1 + \frac{kT}{M^2 \overline{\omega^4}} \left[\frac{1}{3M\overline{\omega^2}} \sum_{\substack{mn \\ ijk}} \Phi^{mn0^2}_{ijk} - \frac{1}{2} \sum_{\substack{m \\ ij}} \Phi^{mm00}_{iijj} \right] \right\}.$$

Für hohe Temperaturen haben wir ein lineares

* In diesen Gliedern kann natürlich wieder $\omega_s(\boldsymbol{k})$ durch $\bar{\omega}_s(\boldsymbol{k})$ ersetzt werden. Vgl. Fussnote Seite 288 wegen der Abweichungen.

VOL. 4 1958

Ansteigen (oder Abfallen) der spezifischen Wärme über den klassischen Wert hinaus. Das kann man auch schon klassisch berechnen.[13,14] Etwas modifiziert wird auch das Einmünden in den klassischen Wert; effektiv kommen jetzt alle negativen Potenzen von T vor ($\overline{\omega^2}$ hängt über die Parameter A von T ab, doch kann in Produkten von $\overline{\omega^2}$ mit Φ_3 oder Φ_4, $A = \bar{A}$ gesetzt werden.), wobei natürlich die geraden Potenzen der harmonischen Theorie viel grössere Beiträge liefern als die ungeraden der anharmonischen Glieder.

Für unser einfaches Modell (Zentralkräfte nur zwischen nächsten Nachbarn) erhalten wir dann leicht

$$C_v/3Nk = 1+\frac{kT}{32f^2}[g^2/f-4h+6\sqrt{2}g/a-30f/a^2]- \tag{4.6a}$$

$$-\frac{1}{3}\cdot\frac{\hbar^2 f}{M(kT)^2}\left\{1-\frac{kT}{4f^2}[g^2/f+4\sqrt{2}g/a+8f/a^2]\right\}+$$

$$+\frac{1}{12}\frac{\hbar^4 f^2}{M^2(kT)^4}\times$$

$$\times\left\{1-\frac{7kT}{20f^2}[g^2/f+12/7h+12\sqrt{2}g/a+4f/a^2]\right\} \tag{4.6b}$$

und

$$\frac{C_p-C_v}{3Nk} = \frac{kT}{8f^3}(g+2\sqrt{2}f/a)^2\times$$

$$\times\left\{1-\frac{2\hbar^2 f}{3M(kT)^2}+\frac{5}{18}\frac{\hbar^4 f^2}{M^2(kT)^4}\right\}.$$

In Abb. 5 ist der Verlauf von C_p für Krypton aufgetragen und mit den Messungen von CLUSIUS[15] verglichen. Die Übereinstimmung ist nur qualitativ, was wegen der Einfachheit unseres Modells nicht verwunderlich ist. Da die Ausdrücke g^2/f und h gleiches Vorzeichen haben, in (4.6) aber im wesentlichen die Differenzen eingehen, machen sich kleine Ungenauigkeiten (Vernachlässigung höherer Nachbarn, Abweichungen im Potentialverlauf) schon stark bemerkbar.

(b) *Tiefe Temperaturen*

Für den Grenzfall tiefer Temperaturen sind die

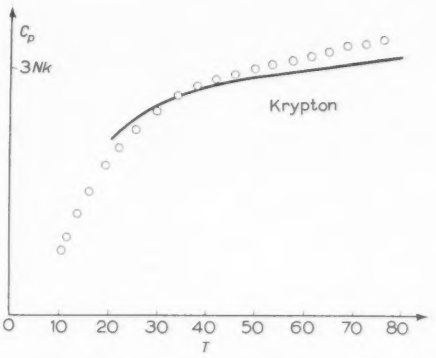


ABB. 5. Spezifische Wärme bei konstantem Druck (C_p) für Krypton.

Kreise : experimentelle Werte.[15]
ausgezogen : berechnet nach dem im Text angegebenen Modell.

Ausdrücke noch etwas unhandlicher. Sie sind für unsere spätere Diskussion aber besonders wichtig. Da alle Nullpunktsanteile in der freien Energie keinen Beitrag zur spezifischen Wärme liefern, seien sie von vornherein unterdrückt. Ferner wollen wir uns auf so tiefe Temperaturen beschränken, dass wir mit der niedrigsten Potenz (d.i. T^4) auskommen.

Der von F_s herrührende Anteil liefert in bekannter Weise :

$$\overset{\hbar}{C}_v/3Nk = \frac{2\pi^2}{5}\cdot\frac{V_z}{\hbar^3}\cdot\overline{\frac{1}{v^3}}(kT)^3 = \frac{4\pi^4}{5}(T/\Theta_v)^3, \tag{4.7}$$

wenn wir die bei tiefen Temperaturen ($T\to 0$) üblichen Näherungen machen : Wir ersetzen die Summe durch ein Integral und vernachlässigen die Gitterstruktur, d.h. wir approximieren das Spektrum durch $z(\omega)\,d\omega \sim \omega^2\,d\omega$ und erstrecken das Integral von 0 bis ∞. Ausserdem verabreden wir folgende Abkürzung : Die Grösse v soll formal so gebildet werden wie in der harmonischen Näherung die Schallgeschwindigkeiten, nur dass an die Stelle der harmonischen Kopplungsparameter $\tilde{\Phi}^{mn}_{ij}$ die Φ^{mn}_{ij} treten. Es wird sich nachher herausstellen, dass die v nicht die Bedeutung von Schallgeschwindigkeiten haben, sondern nur im Grenzfall verschwindender Anharmonizität in diese übergehen! Entsprechend sind die Debyesche Abschneidefrequenz ω_v und die Debye-Temperatur Θ_v definiert :

$$\frac{1}{\omega_v^3} = \frac{V_z}{6\pi^2} \cdot \frac{\overline{1}}{v^3}\,; \quad \hbar\omega_v = k\Theta_v. \tag{4.8}$$

Bei der Auswertung der Gleichungen (2.10b) für tiefe Temperaturen können wir alle Glieder vernachlässigen, die Produkte von zwei mittleren Besetzungszahlen (z.B. $\overline{n_s^{\boldsymbol{k}}} \cdot \overline{n_{s'}^{\boldsymbol{k}'}}$) enthalten. Sie liefern Beiträge, die zu höheren Potenzen der Temperatur als T^4 proportional sind. Denn wir können dann sowohl die Integration über ω als auch über ω' von 0 bis ∞ erstrecken. Die Substitution $\hbar\omega/kT = \eta$; $\hbar\omega'/kT = \eta'$ liefert für jedes ω und ω' einen Faktor T. Zählen wir alle Potenzen von ω und ω' ab, so finden wir die Behauptung bestätigt. Bei den in den Besetzungszahlen $\overline{n_s^{\boldsymbol{k}}}$ linearen Termen können wir nur die Integration über ω von 0 bis ∞ erstrecken, die über ω' läuft bis zum Ende des Spektrums. Das liefert Glieder $\sim T^4$. Aus (2.10b) wird

$$F_3 = -\frac{\hbar^2}{2N} \sum_{\substack{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}'' \\ s\,s'\,s''}} \omega\omega'\omega''\delta(\boldsymbol{k}+\boldsymbol{k}'+\boldsymbol{k}''-2\pi\boldsymbol{b})\times$$

$$\times P_{ss's''}^{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}''} \left(\frac{\tilde{n}}{\omega+\omega'+\omega''} + \frac{\overline{n''}}{\omega+\omega'-\omega''}\right) \tag{4.9}$$

$$F_4 = \frac{\hbar^2}{2N} \sum_{\substack{\boldsymbol{k}\boldsymbol{k}' \\ s\,s'}} \omega\omega' Q_{ss'}^{\boldsymbol{k}\boldsymbol{k}'} \cdot \tilde{n}$$

mit

$$\delta(\boldsymbol{k}+\boldsymbol{k}'+\boldsymbol{k}''-2\pi\boldsymbol{b})P_{ss's''}^{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}''} = \frac{|\Phi_{s\,s'\,s''}^{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}''}|^2}{8\omega^2\omega'^2\omega''^2}\,; \tag{4.10}$$

$$Q_{ss'}^{\boldsymbol{k}\boldsymbol{k}'} = \frac{\Phi_{s\,s\,s'\,s'}^{\boldsymbol{k}-\boldsymbol{k}\boldsymbol{k}'-\boldsymbol{k}'}}{4\omega^2\omega'^2}.$$

$\delta(\boldsymbol{k}+\boldsymbol{k}'+\boldsymbol{k}''-2\pi\boldsymbol{b})$ ist ein Kroneckersymbol, das den Pseudoimpulssatz für die Schallquanten enthält. Es ist nur dann von Null verschieden, wenn die Summe $\boldsymbol{k}+\boldsymbol{k}'+\boldsymbol{k}''$ gleich dem 2π-fachen eines Vektors $\boldsymbol{b}$ des reziproken Gitters ist. $P_{ss's''}^{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}''} = P$ und $Q_{ss'}^{\boldsymbol{k}\boldsymbol{k}'} = Q$ können (für lange Wellen) wir in Näherung als unabhängig von $\boldsymbol{k}, \ldots$ ansehen.

Die Ausführung der Summen liefert (nach Übergang zum Integral)*:

$$F_3+F_4 = -\frac{\pi^2}{10} \cdot \frac{3}{2} N\hbar\overline{\omega} V_z(3P-Q)\frac{\overline{1}}{v^3}\frac{(kT)^4}{\hbar^3} \tag{4.11}$$

$\frac{3}{2}N\hbar\overline{\omega} = U_h(0)$ ist die Nullpunktsenergie. Die spezifische Wärme ist also

$$C_v/3Nk = \frac{4\pi^4}{5}(T/\Theta_v)^3\times$$

$$\times\{1+3(3P-Q)\cdot U_h(0)/N\} = \frac{4\pi^4}{5}(T/\Theta_0)^3. \tag{4.12}$$

Die Debye-Temperatur für $T \to 0$ ist demnach

$$\Theta(0) = \Theta_0 = \Theta_v\{1-(3P-Q)U_h(0)/N\} \tag{4.13}$$

Wegen unserer Definition von v und Θ_v ist es klar, dass diese Grössen durch das Θ_D der harmonischen Näherung ausgedrückt werden können, und zwar können wir für $T \to 0$ den Zusammenhang in der Form†

$$\Theta_v = \Theta_D\{1-P'\cdot U_h(0)/N\} \tag{4.14}$$

schreiben, so dass

$$\Theta_0 = \Theta_D\{1-(3P+P'-Q)U_h(0)/N\}. \tag{4.15}$$

Θ_D kann durch die elastischen Konstanten der harmonischen Näherung ausgedrückt werden, also durch die von hohen Temperaturen linear auf $T = 0$ extrapolierten Konstanten.‡

---

* Hier könnten für v die Werte der harmonischen Näherung benutzt werden. Vgl. weiter oben.

† Für P' ergibt sich.

$$P'U_h(0)/N = -\frac{1}{2\sum_i \Phi_{ii}^{oo}} \sum_{\substack{o \\ ikl}} \Phi_{iik}^{ooo}(Ao)_i V_k^{\text{th}}(0)$$

$V_{(kl)}^{(\text{th})}(0)$ ist aus Gl. (3.4a) oder bei kubischer Symmetrie aus Gl. (3.5) zu entnehmen.

‡ Vergleiche dazu das Diagramm in Literatur (1), S.252, das auf einer Methode von QUIMBEY und SUTTON[16] beruht. Man kann mit seiner Hilfe die Debye-Temperatur der harmonischen Näherung durch die elastischen Konstanten (der harmonischen Näherung) ausdrücken.

VOL. 4 1958

Für unser einfaches Modell ergibt sich

$$\Theta_0 = \Theta_D \times$$

$$\times \{1 - \frac{1}{8f^2}(g^2/f - \tfrac{1}{2}h + 10\sqrt{2}g/a + 48f/a^2) U_h(0)/N\}. \tag{4.16}$$

Die Abweichung vom Θ_D-Wert der harmonischen Theorie ist danach für Krypton etwa 4 % und zwar ist der richtige Wert kleiner als Θ_D. Natürlich ist dieser Unterschied bei den Edelgasen besonders gross; bei anderen Bindungstypen dürfen wir nur kleinere Werte erwarten. Ausserdem hängt er von der Masse M ab, und zwar sind die Abweichungen um so grösser, je kleiner die Masse ist. Auf den Zusammenhang zwischen Debye-Temperatur und den richtigen elastischen Konstanten kommen wir unten in der Diskussion zurück.

5. DIE SCHALLGESCHWINDIGKEIT

Die Schallausbreitung geschieht durch eine Überlagerung von mechanisch angeregten Schwingungen über die immer vorhandenen thermischen Bewegungen. Die Wellenlängen der Schallwellen sind gross gegen die Gitterkonstante der Kristalle. In der harmonischen Theorie erweisen sich die Eigenfrequenzen des Gitters als massgebend für die Schallgeschwindigkeit. Es sind die langen elastischen Wellen, für die die Frequenz ω der Wellenzahl $\boldsymbol{k}$ proportional ist, d.h. Phasengeschwindigkeit ω/k und Gruppengeschwindigkeit $d\omega/dk$ sind gleich. Die Amplituden der mechanisch angeregten und der thermischen Schwingungen addieren sich. Die Anharmonizität bewirkt eine Wechselwirkung zwischen den Schallwellen und thermischen Bewegungen, deren Einfluss genauer untersucht werden muss. Es wird sich herausstellen, dass es in dem Ausdruck für die Schallgeschwindigkeit auf das Verhältnis der Amplituden von mechanischen und thermischen Schwingungen ankommt. Man kann sich leicht klar machen, dass selbst am absoluten Nullpunkt die "thermischen" Amplituden noch zehnmal grösser sind als die mechanischen, die man bei maximaler Belastung erreichen kann.

Zur Vereinfachung untersuchen wir die Verhältnisse an der linearen Kette, bei der sich alles etwas übersichtlicher darstellen lässt. Die Verallgemeinerung auf den räumlichen Fall lässt sich ohne Schwierigkeiten bewerkstelligen. Wir betrachten eine lineare Kette, die an einem Ende periodisch erregt wird mit der Frequenz Ω. Im stationären Fall bilden sich stehende Wellen aus, deren Wellenlänge $1/k$ wir messen können. Aus Ω/k lässt sich die Phasengeschwindigkeit bestimmen. Darüber überlagert sind natürlich immer noch die thermischen Schwingungen. Um die Rechnung zu vereinfachen, betrachten wir nicht eine endliche Kette, sondern eine unendliche, die wir in Periodizitätsvolumina der Länge Na einteilen.* Die erregende Kraft K_0 müssen wir dann auch periodisch angreifen lassen, etwa an den Atomen ... 0, N, $2N$, ... Die Bewegungsgleichungen lauten nach der üblichen Transformation auf Normalkoordinaten (a^k) :

$$-\ddot{a}^k = \omega_k^2 a^k + \sum_{k',k''} g^{kk'k''} a^{k'} a^{k''} +$$

$$+ \sum_{k',k'',k'''} h^{kk'k''k'''} a^{k'} a^{k''} a^{k'''} - \frac{K_0}{2M\sqrt{N}} \times$$

$$\times [\exp(i\Omega t) + \exp(-i\Omega t)] \tag{5.1}$$

mit†

$$\omega_k^2 = \frac{4f}{M} \sin^2(ka/2)$$

$$g^{kk'k''} = -\frac{i4g}{M\sqrt{N}} \delta(k - k' - k'') \times$$

$$\times \sin(ka/2) \sin(k'a/2) \sin(k''a/2) \tag{5.1a}$$

$$h^{kk'k''k'''} = -\frac{8h}{MN} \delta(k - k' - k'' - k''') \times$$

$$\times \sin(ka/2) \sin(k'a/2) \sin(k''a/2) \sin(k'''a/2)$$

---

* Die Rechnung lässt sich auch für die endliche Kette durchführen und liefert praktisch dasselbe Ergebnis. Nur die Durchführung ist etwas umfangreicher.

† f,g,h sind die Konstanten in der potentiellen Energie für die lineare Kette, die die Form hat :

$$\Phi = \Phi_0 + \sum_n \{\tfrac{1}{2} f(s_n - s_{n-1})^2 + \tfrac{1}{3!} g(s_n - s_{n-1})^3 + \tfrac{1}{4!} h(s_n - s_{n-1})^4 + \ldots\}.$$

s_n sind die Auslenkungen der Atome aus ihren mittleren Lagen.

Die Gleichung (5.1) lässt sich im stationären Fall mit dem bei allen erzwungenen Schwingungsproblemen gebräuchlichen Ansatz lösen :

$$a^k = \sum_{\rho=-\infty}^{+\infty} [c_\rho^k \exp(i\nu\Omega t) + d_\rho^k \exp(i\nu\omega_k t)]. \tag{5.2}$$

Er trägt der Überlagerung von mechanisch angeregten (c_ρ^k) und thermischen (d_ρ^k) Schwingungen Rechnung. Da wir immer nur Glieder berücksichtigen, die keine höheren Potenzen der Anharmonizität als g^2 und h enthalten, können wir die Summe (5.2) an einer geeigneten Stelle abbrechen. Mit (5.2) erhalten wir aus (5.1) ein algebraisches Gleichungssystem, das sich rekursiv lösen lässt. Bei der Lösung machen wir Gebrauch von der eingangs erwähnten Tatsache, dass die mechanischen Amplituden klein sind gegen die thermischen (maximal 10 %). Wir können deshalb dort, wo mechanische Amplituden in Konkurrenz zu thermischen auftreten, die ersteren vernachlässigen. Sodann ersetzen wir in einer weiteren Näherung die thermischen Amplituden durch ihre Mittelwerte, die nur für quadratische Ausdrücke von Null verschieden sind. In der harmonischen Näherung gilt†

$$\overline{a^{*k}_{\ 1}\, a^{k'}_{\ 1}} = \overline{d^{*k}_{\ 1}\, d^{k'}_{\ 1}} + \overline{d^{*k}_{-1}\, d^{k'}_{-1}}$$

$$= \frac{\epsilon(k)}{M\omega_k^2}\delta_{kk'};\ \epsilon(k) = \hbar\omega_k(\bar{n}_k + \tfrac{1}{2}). \tag{5.3}$$

Dabei müssen wir die Bedingungen $c^{*k}_\rho = c^{-k}_{-\rho}$; $d^{*k}_\rho = d^{-k}_{-\rho}$ beachten, die aus der Realitätsbedingung $a^{*k} = a^{-k}$ für die lineare Kette folgen (die Auslenkungen s_n sind reell!) Schliesslich erhalten wir mit $\sum_{k''} \epsilon(k'') = U_h$

$$\Omega^2\, c^{k}_{\pm 1} = \omega^2_{k\pm 1}\, c^{k}_{\pm 1} \cdot \left\{1 + \frac{h}{2f^2}\cdot\frac{U_h}{N}\right\} - \frac{K_0}{2M\sqrt{N}}. \tag{5.4}$$

Dieser Ausdruck unterscheidet sich von dem der harmonischen Näherung nur dadurch, dass die geschweifte Klammer als Faktor bei ω_k^2 auftritt. Der Zusammenhang zwischen Frequenz und Wellenlänge ist also mit diesem Faktor zu multiplizieren, wenn wir die Anharmonizität berücksichtigen. Für lange Wellen ist auch hier die Frequenz proportional zur Wellenzahl; die Phasengeschwindigkeit ist unabhängig von der Frequenz. Unsere Rechnung bezieht sich allerdings nur auf die Phasengeschwindigkeit; um Aussagen über die Gruppengeschwindigkeit zu erhalten, müssen wir die Ausbreitung von Impulsen in der Kette studieren. Da sich aber im vorstehenden der Zusammenhang zwischen Frequenz und Wellenzahl als linear erwiesen hat (für lange Wellen), dürfen wir annehmen, dass für die Gruppengeschwindigkeit dieselbe Beziehung gilt.

Die Verallgemeinerung auf den räumlichen Fall ist nicht schwierig. Es ergibt sich

$$\Omega^2 = \omega_s^2(\boldsymbol{k}) + \frac{1}{2MN}\sum_{\boldsymbol{k}'s'} \frac{\Phi^{\boldsymbol{k}-\boldsymbol{k}\boldsymbol{k}'-\boldsymbol{k}'}_{s\ \ s\ s'\ s'}}{M\omega^2_{s'}(\boldsymbol{k}')}\,\epsilon_{s'}(\boldsymbol{k}') \tag{5.5}$$

$\omega_s^2(\boldsymbol{k})$ und $\Phi^{\boldsymbol{k}-\boldsymbol{k}\boldsymbol{k}'-\boldsymbol{k}'}_{s\ \ s\ s'\ s'}$ sind für kleine $\boldsymbol{k}$ (lange Wellen) proportional zu k^2, also ist Ω proportional zu k; Ω/k und $d\Omega/dk$ sind unabhängig von k. Gl. (5.5) lässt sich näherungsweise noch vereinfachen; wir ziehen $\Phi^{\boldsymbol{k}-\boldsymbol{k}\boldsymbol{k}'-\boldsymbol{k}'}_{s\ \ s\ s'\ s'}$ und $\omega^2_{s'}(\boldsymbol{k}')$ als Mittelwerte vor das Summenzeichen, und zwar sollen Zähler und Nenner unabhängig über $\boldsymbol{k}'$ und s' gemittelt werden (der Fehler ist nicht gross, da sich $\boldsymbol{k}'$- und s'- Abhängigkeit in Zähler und Nenner annähernd kompensieren). Da $\sum_{\boldsymbol{k}'s'} \epsilon_{s'}(\boldsymbol{k}') = U_h$, erhalten wir

$$\Omega^2 = \omega_s^2(\boldsymbol{k}) + \frac{\overline{\Phi^{\boldsymbol{k}-\boldsymbol{k}\boldsymbol{k}'-\boldsymbol{k}'}_{s\ \ s\ s'\ s'}}^{\boldsymbol{k}'s'}}{2M^2\overline{\omega^2}}\cdot\frac{U_h}{N}. \tag{5.5a}$$

Für unser spezielles Modell (Federbindungen zu nächsten Nachbarn) ergibt sich nach Auswertung dieser Gleichung, dass sich die Schallgeschwindigkeiten durch eine effektive Federkonstante

$$\hat{f} = f\left\{1 - \frac{1}{12f^2}(g^2/f - h - \sqrt{2}g/a + 4f/a^2)\,U_h/N\right\} \tag{5.6}$$

beschreiben lassen, unabhängig von der Richtung des Ausbreitungsvektors $\boldsymbol{k}$ und der Polarisation s.

Betrachten wir eine Wellenausbreitung in Richtung einer Würfelkante, also $\boldsymbol{k} = (k, 0, 0)$ so folgt dafür

$$\rho c_l^2 = \frac{2\hat{f}}{a};\ \rho c_t^2 = \frac{\hat{f}}{a} \tag{5.7}$$

† Vgl. etwa HUSIMI.[17]

VOL. 4 1958

Nach der elastischen Theorie sollte nun $\rho c_l^2 = c_{11}(T)$; $\rho c_t^2 = c_{44}(T)$ sein. Ein Vergleich mit (3.11) lehrt uns, dass die elastischen Beziehungen wohl für transversale Schallwellen erfüllt sind, nicht dagegen für die longitudinalen. Genau derselbe Sachverhalt ergibt sich für andere Ausbreitungsrichtungen; so sollte in der Flächendiagonalen $\boldsymbol{k} = (1/\sqrt{2})(k, k, 0)$ z.B. gelten

$$\rho c_l^2 = \frac{c_{11}+c_{12}+2c_{44}}{2};$$

$$\rho c_t^2 = c_{44} \text{ und } \rho c_t^2 = \frac{c_{11}-c_{12}}{2}$$

und in der Raumdiagonalen $\boldsymbol{k} = 1/\sqrt{3}(k, k, k)$

$$\rho c_l^2 = \frac{c_{11}+2c_{12}+4c_{44}}{3}; \quad \rho c_t^2 = \frac{c_{11}-c_{12}+c_{44}}{3}.$$

Auch hier gelten nach wie vor die harmonischen Beziehungen für die Transversalwellen, während sich bei den Longitudinalwellen Abweichungen bemerkbar machen. Dass Unterschiede für die beiden Wellentypen vorhanden sind, ist wohl nicht so sehr erstaunlich, denn die Beanspruchung der Federn ist für longitudinale Wellen eine ganz andere als für transversale. Für transversale Ausbreitungen stimmen ja auch allgemein die betreffenden Ausdrück überein, ganz gleich, ob sie mit adiabatischen oder isothermen elastischen Konstanten gebildet werden, während bei den longitudinalen Unterschiede vorhanden sind. Zu bemerken ist dabei, dass die Ausdrücke für die longitudinalen Geschwindigkeiten weder mit den adiabatischen noch mit den isothermen elastischen Konstanten gebildet werden können.

Den Ausdruck (5.5) hat HOOTON[9] für die lineare Kette schon auf ganz andere Weise abgeleitet. Er schreibt den Hamilton-Operator in der Form:

$$\mathscr{H} = \underbrace{\sum_{k,s} \tfrac{1}{2}(\dot{a}^*{}_s^k \dot{a}_s^k + \Omega_s^2(k) a^*{}_s^k a_s^k) +}_{\mathscr{H}_0}$$

$$+ \underbrace{\sum_{k,s} \tfrac{1}{2}(\omega_s^2(k) - \Omega_s^2(k)) a^*{}_s^k a_s^k + \text{anharm. Glieder}}_{W} \tag{5.8}$$

und bestimmt die Frequenzen $\Omega_s(k)$ aus der Forderung

$$\partial F/\partial \Omega_s(k) = 0.$$

Aus (5.8) ist aber ersichtlich, dass der Hamilton-Operator und damit die freie Energie gar nicht von $\Omega_s(k)$ abhängen können. Hooton erhält deswegen einen Ausdruck für $\Omega_s(k)$, weil er nach der Störung W entwickelt; diese Entwicklung bricht er nach dem ersten Glied ab. Dann hängt die freie Energie natürlich effektiv von $\Omega_s(k)$ ab.

Wir wollen das Hootonsche Verfahren noch kurz von einer anderen Seite her beleuchten. Auf der einen Seite betrachten wir die von uns berechnete freie Energie (2.10). Sie sei mit $F_R(\omega_J)$ bezeichnet. Andererseits führen wir ein System von unabhängigen harmonischen Oszillatoren mit unbekannten Frequenzen Ω_J ein; dessen freie Energie ist

$$F_H(\Omega_J) = kT \cdot \sum_J \ln[2 \sinh \tfrac{1}{2}\beta\hbar\Omega_J].$$

Die Ω_J sollen so bestimmt werden, dass das anharmonische System (F_R) durch das harmonische (F_H) möglichst gut dargestellt wird, d.h. für alle physikalischen Grössen sich möglichst dieselben Werte ergeben, ganz gleich aus welcher freien Energie (F_R oder F_H) wir sie bestimmen. Die physikalischen Grössen sind Ableitungen der freien Energie; wir müssen also verlangen, dass die Ableitungen nach Volumen und Temperatur übereinstimmen. Da es $3N$ Frequenzen gibt, müssen wir natürlich ebensoviele Ableitungen heranziehen, was aber sehr kompliziert ist.

Wir können diese Forderungen jedoch durch andere ersetzen, von denen sich zeigen lässt, dass sie im Rahmen unserer Näherung dasselbe liefern. Dazu werden dem harmonischen Ersatzsystem Koordinaten A_J zugeordnet, die temperaturabhängig sind. Wir verlangen Gleichheit der Absolutquadrate im thermischen Mittel mit denen unserer vorhergehenden Rechnung, d.h.

$$\overline{A_J^* A_J} = \overline{a_J^* a_J}$$

für alle J. Die Rechnung wollen wir unterdrücken. Für die Frequenzen Ω_J erhalten wir den Hootonschen Ausdruck, wenn wir *nur* F_4 berücksichtigen; zusätzlich gibt es noch einen Beitrag von F_3, der bei HOOTON nicht auftreten kann, da er seine Entwicklung nach der ersten Näherung abbricht.

Ob sich den Ω_J aber irgendeine physikalische Bedeutung beimessen lässt, ist hieraus nicht ersichtlich. Sie sind nur so gewählt, dass sich die physikalisch messbaren Grössen durch die Ableitungen der freien Energie eines Systems harmonischer Oszillatoren darstellen lassen.

6. DISKUSSION UND VERGLEICH MIT DEM EXPERIMENT

In der harmonischen Theorie lassen sich zwischen den elastischen Konstanten, der Schallgeschwindigkeit und der Debye-Temperatur eindeutige Beziehungen herstellen, aus denen die atomistischen Kopplungsgrössen eliminiert sind.

Unser Ziel ist es, möglichst auch für den Fall, dass wir die Anharmonizität in die Theorie einbeziehen, entsprechende Zusammenhänge aufzuzeigen.

In den vorhergehenden Abschnitten haben wir, in sehr allgemeiner Form, zunächst die Relationen zwischen den makroskopischen Eigenschaften der Kristalle und den atomaren Kopplungsgrössen hergeleitet. Wir müssen dabei zwischen zwei Gruppen unterscheiden :

(1) Die Grössen, die durch die thermische Zustandsgleichung bestimmt sind, wie thermische Ausdehnung und elastische Konstanten.

(2) Die Grössen, die in die kalorische Zustandsgleichung eingehen, wie innere Energie und spezifische Wärmen.*

liegenden Messungen (an Cu, Au, Si, Ge, Pb, Li F, KCl, KBr, MgO) sehr gut bestätigt† (vgl. auch die Rechnungen an Ionenkristallen[18].)

Es lassen sich aber noch einige weitere Aussagen machen. Wir hatten schon darauf hingewiesen (vgl. Abschnitt 3, vorletzter Absatz), dass in kubischen Kristallen, wenn nur Zentralkräfte wirksam sind, immer $c_{12} \ll c_{44}$ sein sollte; beide Konstanten unterscheiden sich nur um quadratische Glieder *eines* Vorzeichens. Vergleichen wir das mit dem Experiment, so finden wir z.B.

$c_{12} < c_{44}$ bei : C, Si, Ge ; LiF, NaCl, KCl, MgO

$c_{12} > c_{44}$ bei : Cu, Ag, Au ; Al, Pb, W; Fe, Ni.

Bei der letzten Gruppe müssen also sehr starke

VOL. 4 1958

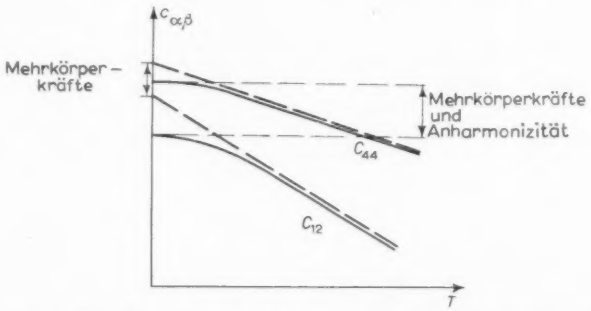


ABB. 6. Zum Einfluss von Mehrkörperkräften und Anharmonizität auf die Cauchy-Relation.

In die Eigenschaften der ersten Gruppe geht die Anharmonizität nur über die eingeführten Parameter, die mittleren Lagen, ein, während wir für die zweite Gruppe höhere Entwicklungsglieder der freien Energie berücksichtigen müssen. Letztlich ist der Einfluss der Anharmonizität aber für thermische und kalorische Eigenschaften gleich gross.

Die Temperaturabhängigkeit der elastischen Konstanten lässt sich in sehr allgemeiner Weise angeben : Die adiabatischen Konstanten sollten proportional zur inneren Energie sein; die isothermen enthalten ausserdem noch einen Term, der proportional zu dem Produkt aus Temperatur und spezifischer Wärme ist. Das wird durch die vor-

Mehrkörperkräfte vorhanden sein. Dass bei der ersten Gruppe $c_{12} < c_{44}$ ist, bedeutet jedoch keineswegs, dass *nur* Zentralkräfte wirksam sind. Denn das erfordert weiterhin, dass die von hohen Temperaturen linear auf $T = 0$ extrapolierten elastischen Konstanten c_{12} und c_{44} (harmonische Näherung) übereinstimmen müssen. Erfüllt ist das ganz gut für Ionenkristalle (KCl, NaCl, NaBr und LiF); bei anderen Kristallen zeigen sich zum Teil erhebliche Abweichungen.

Auch die Cauchy-Relationen ($c_{12} = c_{44}$ im kubischen Kristall) sind demgemäss nie befriedigt. Selbst wenn keine Mehrkörperkräfte vorhanden wären, würden doch Anharmonizität und Nullpunktsenergie dafür sorgen, dass Abweichungen bestehen. Wir sind aber in der Lage, den Einfluss

* Dazu würde auch die Entropie gehören, auf die wir nicht eingegangen sind. Man kann sie aber jederzeit aus der freien Energie (2.10a, 2.10b) bestimmen.

† Eine Zusammenfassung der Ergebnisse erscheint demnächst im Landolt-Börnstein.

beider Effekte auf die Cauchy-Relationen zu trennen, wie es in Abb. 6 skizziert ist.

Bei der Auswertung der Beziehungen für die spezifische Wärme wurden verschiedene Näherungen gemacht, indem wir Summen über die Wellenzahlvektoren $\boldsymbol{k}$ durch ihre Mittelwerte ersetzten. Der Fehler dürfte gering sein, da die Summen durchweg so beschaffen sind, dass im Zähler und Nenner die $\boldsymbol{k}$-Abhängigkeit von gleicher Art ist. Die Debye-Temperatur haben wir durch die atomaren Kopplungsgrössen ausdrücken können. Uns interessiert dabei speziell die Debye-Temperatur im Grenzfall sehr tiefer Temperaturen ($T \to 0$); nur dann können wir vom Einfluss des Gitters auf das Spektrum absehen und es durch das des elastischen Kontinuums ersetzen. Eine genaue Berücksichtigung des Gitterspektrums würde die Sache ungeheuer kompliziert machen, was wir ausser Betracht lassen wollen.

Jetzt kommen wir auf unsere primäre Frage zurück, ob es möglich ist, ähnlich wie in der harmonischen Theorie, die Debye-Temperatur für $T = 0$ allein durch elastische Daten auszudrücken. Die Beziehung zwischen dem richtigen Wert Θ-Wert am absoluten Nullpunkt (d.h. Θ_0) und dem Θ_D der harmonischen Näherung ist in (4.15) angegeben. Θ_D können wir aus den linear auf $T = 0$ extrapolierten elastischen Konstanten ermitteln, wobei wir die von QUIMBEY und SUTTON[16] angegebenen Relationen (für kubische Kristalle) benutzen :

$$k\Theta_D = \hbar\omega_D = 2\pi\hbar \left(\frac{3}{4\pi V_z}\right)^{\frac{1}{3}} \times \times \left(\frac{c_{11}-c_{44}}{\rho}\right)^{\frac{1}{2}} \cdot J(C,K) \qquad (6.1)$$

mit

$$C = c_{44}/c_{11}-c_{44}\,;\quad K = c_{12}+c_{44}/c_{11}-c_{44}$$

$J(C, K)$ ist eine relativ komplizierte Funktion und z.B. bei Leibfried[1] tabelliert.

Wir wollen nun in Gl. (6.1) versuchsweise die richtigen elastischen Konstanten für $T = 0$ und $T = \Theta_D$ einsetzen und die so resultierenden „Debye-Temperaturen" Θ' und Θ'' mit Θ_0 vergleichen. Wir benutzen die adiabatischen elastischen Konstanten* und erhalten :

$$\Theta' = \Theta_D \left\{1 - \frac{1}{24f^2}(g^2/f - h + 8f/a^2) \cdot k\Theta_D\right\}$$
$$\Theta'' = \Theta_D \left\{1 - \frac{1}{8f^2}(g^2/f - h + 8f/a^2) \cdot k\Theta_D\right\}. \qquad (6.2)$$

Die relative Lage der verschiedenen Θ-Werte ist mit den Daten für Krypton in Abb. 7 dargestellt (für Xenon wären die angegebenen Prozentzahlen um einen Faktor 2/3 kleiner).

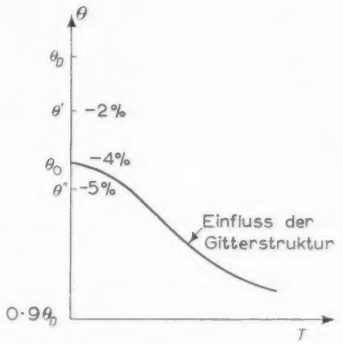


ABB. 7. Relative Lage verschiedener Θ-Werte (vgl. Text). Die angegebenen Prozentzahlen gelten etwa für Krypton.

Leider sind die spezifischen Wärmen und vor allem die elastischen Konstanten der Edelgase bei genügend tiefen Temperaturen nicht gemessen, um quantitative Vergleiche anstellen zu können. Dagegen liegen Messungen vor von CLUSIUS[19] an LiF und von MARTIN[20] an LiF, NaCl und ZnS.

Dürfen wir aber die an unserem Modell gezogenen Schlüsse ohne weiteres auf diese komplizierten Fälle übertragen? Wir dürfen diese Frage für Θ' und Θ'' bejahen ; denn diese Θ-Werte hängen nur vom Gang der elastischen Konstanten mit der Temperatur ab, den wir in allgemeiner Form abgeleitet haben. Natürlich haben die Koeffizienten bei der Temperaturfunktion andere Werte, doch können sie aus dem experimentell ermittelten Verlauf entnommen werden. Danach sollten für die Ionenkristalle die Prozentzahlen in Abb. 7 um die Hälfte kleiner sein.

Etwas schwieriger ist der Vergleich für Θ_0. Es gilt folgendes (vgl. 4.15) : Sicher ist, dass $3P-Q > 0$; wäre $3P-Q < 0$, so müsste die

* In diesem speziellen Modell ergeben sich mit den isothermen elastischen Konstanten dieselben Werte! Ob das in allgemeineren Fällen auch so ist, lässt sich schwer sagen, da $J(C, K)$ nur numerisch bekannt ist.

spezifische Wärme für $T > \Theta$ unter dem klassischen Wert liegen und die elastischen Konstanten $c_{\alpha\beta}$ müssten mit der Temperatur zunehmen, denn in beiden Fällen stehen (bis auf geringfügige Abweichungen in den Zahlenfaktoren) gleiche Ausdrücke vor der Temperaturabhängigkeit. Die Experimente sprechen eindeutig für $3P-Q > 0$. Dieser Anteil würde einen Θ-Wert in der Grösse von Θ' liefern. Dazu kommt noch der Anteil $P' > 0$, der in der Fussnote allgemein angegeben ist. P' ist etwa so gross wie P, d.h. aber $\Theta_0 < \Theta'$; die Differenzen $\Theta_D - \Theta'$ und $\Theta' - \Theta_0$ sollten etwa gleich gross sein.

Beim NaCl ist der Verlauf der elastischen Konstanten einigermassen gut gemessen.[21] Es ergibt sich $\Theta_D = 326°\mathrm{K}$, $\Theta' = 322°\mathrm{K}$; leider schwanken die gemessenen Werte von Θ über einen Bereich von 20° (um 320°K). Beim LiF (723°K gemessen) und ZnS (315°K gemessen) sind die spezifischen Wärmen hinreichend genau gemessen, es fehlen aber Messungen über die Temperaturabhängigkeit der elastischen Konstanten.* Nimmt man an, dass die prozentuale Änderung der elastischen Konstanten mit der Temperatur für alle Ionenkristalle etwa gleich gross ist (soweit Messungen vorliegen, gilt das annähernd), so findet man :

	Θ_D	Θ'	gemessen	Θ''
LiF	750	735	723	717°K
ZnS	330	325	315	306°K

Die gemessenen Werte sollten mit unserem Θ_0 übereinstimmen, das wir nicht angeben können, solange wir die Kopplungsgrössen nicht berechnet haben. Man muss alle aufgeführten Zahlen sehr vorsichtig beurteilen. Zumindest scheinen sich unsere Ergebnisse und die Experimente nicht zu widersprechen. Die spezifische Wärme ist wohl beim LiF am saubersten gemessen. Es wäre gut, auch die elastischen Konstanten im ganzen Temperaturbereich für LiF zu messen.

Von der theoretischen Seite her lassen sich die allgemeinen Formeln ohne prinzipielle Schwierigkeiten auch für andere Fälle auswerten, als wir es hier mit unserem speziellen Modell getan haben (Berücksichtigung von Nachbarn höherer Ordnung, Mehrkörperkräften, Ionenkristallen usw.). Der Rechenaufwand wird unter Umständen allerdings recht erheblich.

Auch der Zusammenhang zwischen elastischen Daten und Schallgeschwindigkeit unterliegt gewissen Änderungen. Bei unserem speziellen Modell hatten wir gesehen, dass die in der harmonischen Theorie abgeleiteten Beziehungen für transversale Wellen (z.B. $\rho c_t^2 = c_{44}(T)$ bzw. $\rho c_t^2 = \frac{1}{2}(c_{11}(T) - c_{12}(T))$ weiterhin gelten. In allen diesen Fällen stimmen die in Frage kommenden Ausdrücke auch für isotherme und adiabatische Konstanten überein. Dagegen ergaben sich Abweichungen für die longitudinalen Schallwellen (z.B. $\rho c_l^2 \neq c_{11}(T)$). Wollen wir hieraus Schlüsse ziehen, müssen wir beachten, dass sich bei allgemeinen Ausbreitungsvektoren $\boldsymbol{k}$ die Einteilung in longitudinale und transversale Schallwellen nicht mehr durchführen lässt (es ist z.B. kein Polarisationsvektor mehr parallel zu $\boldsymbol{k}$). Jedoch lässt sich an Hand unserer Ergebnisse feststellen, dass aus dynamischen Messungen nicht unbedingt die elastischen Konstanten ermittelt werden (z.B. wird die longitudinale Schallgeschwindigkeit in Richtung einer Würfelkante nicht durch c_{11} bestimmt).

Damit lässt sich vielleicht auch RAMANS Behauptung[22] entkräften. Raman hat die aus statischen und dynamischen Messungen resultierenden „elastischen Konstanten" verglichen und meint, systematische Abweichungen gefunden zu haben. Er glaubt, diesen Sachverhalt durch *vier* kubische „elastische Konstanten" darstellen zu können. Dazu muss er aber annehmen, dass der Spannungstensor nicht mehr symmetrisch ist, was unsinnige Folgen nach sich zieht. Wir haben gezeigt, dass Diskrepanzen zwischen statisch und dynamisch ermittelten Konstanten eine einfache Folge der Anharmonizität sind; nach wie vor gibt es nur *drei* elastische Konstanten (im kubischen Fall). Geändert werden nur die Zusammenhänge zwischen elastischen Konstanten und Schallgeschwindigkeit.

Wir sollten vielleicht noch kurz darauf eingehen, wie gut die beiden wesentlichen Näherungen, die wir gemacht haben, erfüllt sind : die Vernachlässigung höherer anharmonischer Glieder (Φ_5, Φ_6, ...) sowie die Behandlung der Anharmonizität (Φ_3, Φ_4) als Störung bei der Berechnung der freien Energie.

Dazu vergleicht man etwa den richtigen Potentialverlauf (z.B. für das als Beispiel herangezogene 6-12-Potential) mit der benutzten Entwicklung bis zu Gliedern vierten Grades an einer Stelle, die der mittleren

* Anmerkung bei der Korrektur: Kürzlich wurden Messungen der elastischen Konstanten von LiF veröffentlicht (C.V. Briscoe, C. F. SQUIRE, *Phys. Rev.* **106,** 1175 (1957)). Wenn man diese benutzt, erhält man genau die in der Tabelle angegebenen Werte für Θ_D und Θ.

VOL. 4 1958

Energie für $T = \Theta$ entspricht (Abb. 8). Bei Edelgasen (ausgenommen Neon) betragen die Differenzen maximal etwa 30 %. Diese Unterschiede würden die höheren Entwicklungsglieder ausgleichen. Bei hohen Temperaturen würden diese höheren Glieder zu einer Temperaturabhängigkeit proportional zu T^2 führen, was experimentell nicht beobachtet ist. Ebenso sind die Abweichungen zwischen quadratischer und unserer Entwicklung bis zu 30 %. Die gemachten Näherungen scheinen deshalb noch vertretbar zu sein. Ferner müssen wir uns daran erinnern, dass die Edelgase wegen der geringen Bindung der ungünstigste Fall sind; bei allen anderen Bindungstypen sind die Potentialmulden tiefer, und alle Näherungen dürften viel besser sein.

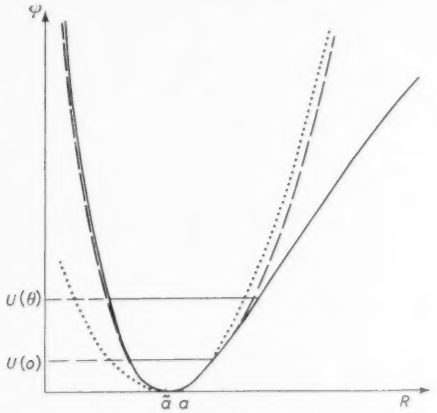


ABB. 8. Zum Vergleich verschiedener Entwicklungen der potentiellen Energie um die mittleren Lagen.

ausgezogen : wahrer Potentialverlauf (6–12-Potential)
gestrichelt : Entwicklung bis zu Gliedern 4. Ordnung einschliesslich
punktiert : Entwicklung bis zu quadratischen Gliedern.

Messungen der thermischen und kalorischen Daten über den ganzen Temperaturbereich, insbesondere auch bei sehr tiefen Temperaturen, liegen leider nur recht wenige vor, so dass die quantitativen Vergleiche sehr spärlich ausgefallen sind. Von der theoretischen Seite her sind die van-der-Waals-Kristalle am einfachsten zu behandeln. Interessant wäre es deshalb, bei ihnen alle thermischen und kalorischen Eigenschaften experimentell bis zu sehr tiefen Temperaturen zu verfolgen und mit der Theorie zu vergleichen. Es liesse sich dann auch der Einfluss aller Näherungsannahmen genauer studieren.

Herrn Prof. Dr. G. LEIBFRIED danke ich für die Anregung zu dieser Arbeit sowie für viele fördernde und klärende Diskussionen.

ANHANG 1

Berechnung von F_3 und F_4

Mit der Störung $W = \Phi_3 + \Phi_4$ sind nach (2.8) und (2.9) F_3 und F_4 zu berechnen. Dazu ist es zweckmässig, durch

$$a_s^{\boldsymbol{k}} = C_s^{\boldsymbol{k}}(b_s^{\boldsymbol{k}} - b^{+-\boldsymbol{k}}_{s}); \quad \left(C_s^{\boldsymbol{k}}\right)^2 = \frac{\hbar}{2\omega_s(\boldsymbol{k})}$$

die bekannten Erzeugungs- und Vernichtungsoperatoren $b^{+\boldsymbol{k}}_{s}$, $b_s^{\boldsymbol{k}}$ für Phononen einzuführen. Sie genügen den Vertauschungsrelationen

$$[b_s^{\boldsymbol{k}}, b^{+\boldsymbol{k}}_{s'}] = \delta_{\boldsymbol{k}\boldsymbol{k}'}\delta_{ss'};$$

$$[b_s^{\boldsymbol{k}}, b_{s'}^{\boldsymbol{k}'}]_- = [b^{+\boldsymbol{k}}_{s}, b^{+\boldsymbol{k}'}_{s'}]_- = 0.$$

Dann ist

$$\Phi_3 = \frac{1}{3!N^{\frac{1}{2}}} \sum_{\substack{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}'' \\ ss's''}} C_s^{\boldsymbol{k}} C_{s'}^{\boldsymbol{k}'} C_{s''}^{\boldsymbol{k}''} \Phi_{ss's''}^{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}''} \times$$

$$\times \{b_s^{\boldsymbol{k}} b_{s'}^{\boldsymbol{k}'} b_{s''}^{\boldsymbol{k}''} - 3b^{+-\boldsymbol{k}}_{s} b_{s'}^{\boldsymbol{k}'} b_{s''}^{\boldsymbol{k}''} + 3b^{+-\boldsymbol{k}}_{s} b^{+-\boldsymbol{k}'}_{s'} b_{s''}^{\boldsymbol{k}''} -$$

$$- b^{+-\boldsymbol{k}}_{s} b^{+-\boldsymbol{k}'}_{s'} b^{+-\boldsymbol{k}''}_{s''} + 3(b^{+-\boldsymbol{k}}_{s} - b_s^{\boldsymbol{k}})\delta_{\boldsymbol{k}',-\boldsymbol{k}''}\delta_{s's''}\}$$

$$\Phi_4 = \frac{1}{4!N} \sum_{\substack{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}''\boldsymbol{k}''' \\ ss's''s'''}} C_s^{\boldsymbol{k}} C_{s'}^{\boldsymbol{k}'} C_{s''}^{\boldsymbol{k}''} C_{s'''}^{\boldsymbol{k}'''} \Phi_{ss's''s'''}^{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{k}''\boldsymbol{k}'''} \times$$

$$\times (b_s^{\boldsymbol{k}} - b^{+-\boldsymbol{k}}_{s})(b_{s'}^{\boldsymbol{k}'} - b^{+-\boldsymbol{k}'}_{s'})(b_{s''}^{\boldsymbol{k}''} - b^{+-\boldsymbol{k}''}_{s''})(b_{s'''}^{\boldsymbol{k}'''} - b^{+-\boldsymbol{k}'''}_{s'''}).$$

Die Diagonalelemente von W sind dann

$$(n = n_s^{\boldsymbol{k}},\ n' = n_{s'}^{\boldsymbol{k}'};\ C = C_s^{\boldsymbol{k}},\ C' = C_{s'}^{\boldsymbol{k}'})^*$$

$$W_{\alpha\alpha} = W_{(n)(n)} = \frac{1}{2N} \sum_{\substack{\boldsymbol{k} \\ s \neq s'}} C^2 C'^2 \Phi_{ss\,s'\,s'}^{\boldsymbol{k}-\boldsymbol{k}\boldsymbol{k}'-\boldsymbol{k}'} \times$$

$$\times (n+\tfrac{1}{2})(n'+\tfrac{1}{2}) + \frac{1}{4N} \sum_{\substack{\boldsymbol{k} \\ s}} C^4 \Phi_{ss\,ss}^{\boldsymbol{k}-\boldsymbol{k}\boldsymbol{k}-\boldsymbol{k}} (n+\tfrac{1}{2})^2.$$

---

* In der Darstellung, in der

$$\mathcal{H}_0 = \Phi_0 + \sum_{\boldsymbol{k}\,s} \hbar\omega_s(\boldsymbol{k})\{b^{+\boldsymbol{k}}_{s} b_s^{\boldsymbol{k}} + \tfrac{1}{2}\}$$

diagonal ist. Die Zustände sind durch $\ldots n_s^{\boldsymbol{k}} \ldots = (n)$ indiziert.

Zu berechnen ist :

$$F_4 = \frac{\sum\limits_{(n)} W_{(n)(n)} \exp(-\beta \sum\limits_{k,s} \hbar\omega n)}{\sum\limits_{(n)} \exp(-\beta \sum\limits_{k,s} \hbar\omega n)} = \overline{W_{(n)(n)}}.$$

Mit

$$\overline{n^2} = 2\bar{n}^2 + \bar{n} \; ; \; \bar{n} = \frac{1}{\exp(\beta\hbar\omega) - 1}$$

lassen sich die beiden Teilsummen über $\overset{k}{s} \neq \overset{k'}{s'}$ und über $\overset{k}{s}$ zusammenfassen und liefern den in (2.10b) angegebenen Ausdruck.

Etwas komplizierter gestaltet sich die Berechnung von F_3, da dort die Nichtdiagonalelemente von Φ_3 eingehen. In $\Phi^{kk'k''}_{ss's''}$ ist ein Kroneckersymbol enthalten, nämlich $\delta(k + k' + k'' - 2\pi b)$.* Für die in b^{k}_{s} und b^{+-k}_{s} linearen Glieder wird daraus wegen $\delta_{k',-k''}\delta_{s's''}$: $\delta(k - 2\pi b)$. Da $-\pi b_x < k_x \leqslant \pi b_x$ usw., kommt nur $k = 0$ in Frage ; das bedeutet eine Schwerpunktsbewegung, die hier nicht interessiert. Die in b^{k}_{s} und b^{+-k}_{s} linearen Glieder geben keinen Beitrag. Ferner geben bei Berechnung der Matrixelemente die Fälle keinen Beitrag, wo alle drei oberen und unteren Indizes gleich sind ; denn wegen $\delta(3k - 2\pi b)$ muss entweder $k = 0$ sein ($b = 0$, Normalprozesse), während sich Glieder mit $b = \pm Bn$ (Umklappprozesse) gegenseitig herausheben. So folgt (mit $\underline{n} = n^{-k}_{s}$) :

$$\Phi_{3(m)(n)} = \frac{1}{6N^{\frac{1}{2}}} \sum_{\substack{k \\ s}} \sum_{\substack{\neq k' \\ s'}} \sum_{\substack{\neq k'' \\ s''}} \sum_{\substack{\neq k \\ s}} CC'C'' \Phi^{kk'k''}_{ss's''} \times$$

$$\times \{\sqrt{n \cdot n' \cdot n''}\delta_{m,n-1}\delta_{m',n'-1}\delta_{m'',n''-1} -$$

$$-3\sqrt{(\underline{n}+1)n'n''}\delta_{\underline{m},\underline{n}+1}\delta_{m',n'-1}\delta_{m'',n''-1} +$$

$$+3\sqrt{(\underline{n}+1)(\underline{n}'+1)n''}\delta_{\underline{m},\underline{n}+1}\delta_{\underline{m}',\underline{n}'+1}\delta_{m'',n''-1} +$$

$$+\sqrt{(\underline{n}+1)(\underline{n}'+1)(\underline{n}''+1)}\delta_{\underline{m},\underline{n}+1}\delta_{\underline{m}',\underline{n}'+1}\delta_{\underline{m}'',\underline{n}''+1}\} \times$$

$$\times \prod_{\substack{K \\ s}} \sum_{\substack{\neq k, k', k'' \\ s, s', s''}} \delta_{m^K_s n^K_s} +$$

$$+\frac{1}{2N^{\frac{1}{2}}} \sum_{\substack{k \\ s}} \sum_{\substack{\neq k' \\ s'}} C^2C' \Phi^{kkk'}_{ss's'} \times$$

$$\times \{\sqrt{n(n-1)}n'\delta_{m,n-2}\delta_{m',n'-1} +$$

$$+\sqrt{(\underline{n}+1)(\underline{n}+2)}n'\delta_{\underline{m},\underline{n}+2}\delta_{m',n'+1} -$$

$$-2\sqrt{(\underline{n}+1)}nn'\delta_{\underline{m},\underline{n}+1}\delta_{m,n+1}\delta_{m',n'-1} +$$

$$+2\sqrt{(\underline{n}+1)n(\underline{n}'+1)}\delta_{\underline{m},\underline{n}+1}\delta_{m,n-1}\delta_{\underline{m}',\underline{n}'+1} -$$

$$-\sqrt{(\underline{n}'+1)n(n-1)}\delta_{m,n-2}\delta_{\underline{m}',\underline{n}'+1} -$$

$$-\sqrt{(\underline{n}+1)(\underline{n}+2)(\underline{n}'+1)}\delta_{\underline{m},\underline{n}+2}\delta_{\underline{m}',\underline{n}'+1}\}.$$

Bei der Bildung des Absolutquadrates von $\Phi_{3(m)(n)}$ erhält man nur etwas aus dem Quadrieren der einzelnen Summanden, da die gemischten Glieder wegen des Faktors $\prod\limits_{\substack{K \\ s}} \delta_{m^K_s n^K_s}$ keinen Beitrag liefern. Es bestehen aber noch verschiedene Möglichkeiten hinsichtlich der Verteilung der $\overset{k}{s}, \overset{k'}{s'}, \overset{k''}{s''}$ des ersten Faktors auf die $\overset{k}{s}, \overset{k'}{s'}, \overset{k''}{s''}$ des zweiten Faktors, was zu verschiedenen Zahlenfaktoren Anlass gibt (3!, 2!, oder 1!). Es folgt :

$$|\Phi_{3(m,n)}|^2 = \frac{1}{6N} \sum_{\substack{k \\ s}} \sum_{\substack{\neq k' \\ s'}} \sum_{\substack{\neq k'' \\ s''}} \sum_{\substack{\neq k \\ s}} C^2C'^2C''^2 |\Phi^{kk'k''}_{ss's''}|^2 \times$$

$$\times \{n \cdot n' \cdot n''\delta_{m,n-1}\delta_{m',n'-1}\delta_{m'',n''-1} +$$

$$+3(\underline{n}+1)n'n''\delta_{\underline{m},\underline{n}+1}\delta_{m',n'-1}\delta_{m'',n''-1} +$$

$$+3(\underline{n}+1)(\underline{n}'+1)n''\delta_{\underline{m},\underline{n}+1}\delta_{\underline{m}',\underline{n}'+1}\delta_{m'',n''+1} +$$

$$+(\underline{n}+1)(\underline{n}'+1)(\underline{n}''+1)\delta_{\underline{m},\underline{n}+1}\delta_{\underline{m}',\underline{n}'+1}\delta_{\underline{m}'',\underline{n}''+1}\} +$$

$$+\frac{1}{4N} \sum_{\substack{k \\ s}} \sum_{\substack{\neq k' \\ s'}} C^4C'^2 |\Phi^{kk'k''}_{ss's''}|^2 \{n(n-1)n'\delta_{m,n-2}\delta_{m',n'-1} +$$

$$+(\underline{n}+1)(\underline{n}+2)n'\delta_{\underline{m},\underline{n}+2}\delta_{m',n'-1} +$$

$$+4(\underline{n}+1)nn'\delta_{\underline{m},\underline{n}+1}\delta_{m,n-1}\delta_{m',n'-1} +$$

$$+4(\underline{n}+1)n(\underline{n}'+1)\delta_{\underline{m},\underline{n}+1}\delta_{m,n-1}\delta_{\underline{m}',\underline{n}'+1} +$$

$$+(\underline{n}'+1)n(n-1)\delta_{m,n-2}\delta_{\underline{m}',\underline{n}'+1} +$$

$$+(\underline{n}+1)(\underline{n}+2)(\underline{n}'+1)\delta_{\underline{m},\underline{n}+2}\delta_{\underline{m}',\underline{n}'+1}\}.$$

Setzen wir

$$\frac{\exp[\beta(E_\alpha - E_\gamma) - 1 - \beta(E_\alpha - E_\gamma)}{(E_\alpha - E_\gamma)^2} = K\left(\frac{E_\alpha - E_\gamma}{\hbar}\right),$$

so wird

$$F_3 = -\frac{1}{\beta} \cdot \left\{ \sum_{(n)} \exp(-\beta \sum_{k,s} \hbar\omega n) \right\}^{-1} \times$$

$$\times \sum_{m,n} |\Phi_{3(m)(n)}|^2 \cdot K(\sum_{k,s} \hbar\omega(m-n)) \times$$

$$\times \exp(-\beta \sum_{k,s} \hbar\omega n).$$

Die Summen lassen sich ausführen und da mit

$$\omega_s(k) = \omega_s(-k) \qquad \bar{n} = \bar{n})$$

ist, zusammenfassen zu:

* b ist ein Vektor des reziproken Gitters.

VOL. 4 1958

$$F_3 = -\frac{1}{6\beta N}\sum_{\substack{k,k',k'\\ s,s',s''}} C^2C'^2C''^2|\Phi_{ss's''}^{kk'k''}|^2\times$$

$$\times\{\bar{n}\cdot\bar{n}'\cdot\bar{n}''\cdot K(\omega+\omega'+\omega'')+$$
$$+3(\bar{n}+1)\cdot\bar{n}'\cdot\bar{n}''\cdot K(-\omega+\omega'+\omega'')+$$
$$+3(\bar{n}+1)(\bar{n}+1)\bar{n}''\cdot K(-\omega-\omega'+\omega'')+$$
$$+(\bar{n}+1)(\bar{n}'+1)(\bar{n}''+1)\cdot K(-\omega-\omega'-\omega'')\}.$$

Das erste und vierte Glied dieser Summe lassen sich auf Grund der Eigenschaften von $K(\omega+\omega'+\omega'')$ vereinigen zu

$$\beta\cdot\frac{(\bar{n}+1)(\bar{n}'+1)(\bar{n}''+1)-\bar{n}\cdot\bar{n}'\cdot\bar{n}''}{\hbar(\omega+\omega'+\omega'')}$$

und ebenso das zweite und dritte Glied zu

$$3\beta\cdot\frac{(\bar{n}+1)(\bar{n}'+1)\bar{n}''-\bar{n}\cdot\bar{n}'\cdot(\bar{n}''+1)}{\hbar(\omega+\omega'-\omega'')}.$$

Vertauscht man an einigen Stellen noch die Summationsbezeichnungen (z.B. $\frac{k}{s}$ mit $\frac{k'}{s'}$), so kann man diese Terme noch vereinfachen zu dem in (2.10b) angegebenen Ausdruck für F_3.

ANHANG 2

Berechnung der thermischen Dehnung V_{ij}^{th} und der elastischen Konstanten

Aus (2.10) erhalten wir für (3.3)

$$\frac{\partial F}{\partial V_{ij}} = \frac{\partial\Phi(Am)}{\partial V_{ij}} + kT\cdot\sum_{k,s}\frac{\partial\omega_s(\boldsymbol{k})}{\partial V_{ij}}\times$$
$$\times\frac{\partial}{\partial\omega_s(\boldsymbol{k})}\ln[2\sinh\tfrac{1}{2}\beta\hbar\omega_s(\boldsymbol{k})] = 0. \quad \text{(A.1)}$$

Entwicklung von $\Phi(Am)$ um die Ruhelagen $\tilde{A}m$ liefert

$$\Phi(Am)=\Phi(\tilde{A}m)+\tfrac{1}{2}\sum_{\substack{m,n\\ ij\ kl}}\tilde{\Phi}_{i\,k}^{mn}(\tilde{A}m)_j(\tilde{A}n)_l V_{ij}V_{kl}+ \quad \text{(A.2)}$$
$$+\frac{1}{3!}\sum_{\substack{mno\\ ijklmn}}\tilde{\Phi}_{i\,km}^{mn\,o}(\tilde{A}m)_j(\tilde{A}n)_l(\tilde{A}o)_n V_{ij}V_{kl}V_{mn}+\ldots$$

also

$$\left(\frac{\partial\Phi(Am)}{\partial V_{ij}}\right)_{V_{kl}=V_{kl}^{\text{th}}} = \sum_{\substack{mn\\ k\,l}}\tilde{\Phi}_{i\,k}^{mn}(\tilde{A}m)_j(\tilde{A}n)_l V_{kl}^{\text{th}}$$
$$= NV_z\cdot\sum_{kl}C_{ij,kl}\cdot V_{kl}^{\text{th}} \quad \text{(A.3)}$$

(wegen der Bezeichnungen siehe Abschnitt 3). Der zweite Teil von (A.1) lässt sich offensichtlich umformen in

$$\tfrac{1}{2}\sum_{k,s}\frac{\partial\ln\omega_s^2(\boldsymbol{k})}{\partial V_{ij}}\cdot\frac{\partial}{\partial\beta}\ln[2\sinh\tfrac{1}{2}\beta\hbar\omega_s(\boldsymbol{k})] \quad \text{(A.4)}$$

worin wir $\partial\ln\omega_s^2(k)/\partial V_{ij}$ durch $\partial\ln\overline{\omega^2}/\partial V_{ij}$ ersetzen wollen (vgl. die Diskussion hierzu weiter unten). Da nun

$$\frac{\partial}{\partial\beta}\sum_{k,s}\ln[2\sinh\tfrac{1}{2}\beta\hbar\omega] = U_h, \quad \text{(A.5)}$$

so folgt durch Einsetzen von (A.2) und (A.3) in (A.1) und Auflösen nach V_{kl}^{th} schliesslich der in (3.4a) angegebene Ausdruck.

Um die elastischen Konstanten zu bekommen, haben wir (A.1) noch einmal abzuleiten und an der Stelle $V_{ij}=V_{ij}^{\text{th}}$ zu nehmen. Der erste Anteil liefert

$$\frac{\partial^2\Phi(Am)}{\partial V_{ij}\partial V_{kl}} = NV_zC_{ij,kl}+$$
$$+\sum_{\substack{mno\\ mn}}\tilde{\Phi}_{i\,km}^{mno}(\tilde{A}m)_j(\tilde{A}n)_l(\tilde{A}o)_n V_{mn}^{\text{th}}, \quad \text{(A.6)}$$

der zweite unter Berücksichtigung von (A.4)

$$\tfrac{1}{2}\sum\left\{\frac{\partial^2\ln\omega_s^2(\boldsymbol{k})}{\partial V_{ij}\partial V_{kl}}\cdot\frac{\partial}{\partial\beta}\ln[2\sinh\tfrac{1}{2}\beta\hbar\omega]+\right.$$
$$\left.+\frac{\partial\ln\omega_s^2(\boldsymbol{k})}{\partial V_{ij}}\cdot\frac{\partial}{\partial\beta}\cdot\frac{\partial}{\partial V_{kl}}\ln[2\sinh\tfrac{1}{2}\beta\hbar\omega]\right\},$$

da

$$\frac{\partial}{\partial\beta}\cdot\frac{\partial}{\partial V_{kl}}\ln[2\sinh\tfrac{1}{2}\beta\hbar\omega]$$
$$=\tfrac{1}{2}\frac{\partial\ln\omega_s^2(\boldsymbol{k})}{\partial V_{kl}}\cdot\frac{\partial}{\partial\beta}\beta\frac{\partial}{\partial\beta}\ln[2\sinh\tfrac{1}{2}\beta\hbar\omega]$$
$$=\tfrac{1}{2}\frac{\partial\ln\omega_s^2(\boldsymbol{k})}{\partial V_{kl}}\left\{\frac{\partial}{\partial\beta}\ln[2\sinh\tfrac{1}{2}\beta\hbar\omega]-\right.$$
$$\left.-T\frac{\partial}{\partial T}\cdot\frac{\partial}{\partial\beta}\ln[2\sinh\tfrac{1}{2}\beta\hbar\omega]\right\}.$$

Ziehen wir in diesen Ausdrücken die Ableitungen von $\ln\omega_s^2(\boldsymbol{k})$ nach den V_{kl} wieder als Ableitungen von $\ln\overline{\omega^2}$ vor das Summenzeichen, so folgen schliesslich die Gln. (3.7).

Wir wollen nun abschätzen, welchen Fehler wir bei

der Ersetzung von $\partial \ln \omega_s^2(\boldsymbol{k})/\partial V_{ij}$ durch $\partial \ln \overline{\omega^2}/\partial V_{ij}$ machen. Wir bedienen uns zunächst (in abgeänderter Form) eines Arguments, das schon Born für hohe Temperaturen benutzt hat[1,12], und übertragen es auf tiefe Temperaturen.

Der zur Diskussion stehende Ausdruck ist

$$\frac{1}{3N}\sum_{k,s}\frac{\partial \ln \omega_s^2(\boldsymbol{k})}{\partial V_{ij}}\cdot\frac{\partial}{\partial\beta}\ln[2\sinh\tfrac{1}{2}\beta\hbar\omega]$$

$$=\begin{cases}\dfrac{kT}{3N}\cdot\dfrac{\partial}{\partial V_{ij}}\displaystyle\sum_{k,s}\ln\omega_s^2(\boldsymbol{k})=kT\cdot\dfrac{\partial}{\partial V_{ij}}\overline{\ln\omega^2} & \\ \qquad\qquad\text{für hohe Temperaturen} & \\ \dfrac{\hbar}{3N}\dfrac{\partial}{\partial V_{ij}}\displaystyle\sum\omega_s(\boldsymbol{k})=\frac{\hbar}{2}\frac{\partial}{\partial V_{ij}}\ln\bar{\omega}^2 & \\ \qquad\qquad\text{für } T=0 & \end{cases} \tag{A.7}$$

Die von uns benutzte Näherung liefert an Stelle obiger Ausdrücke

$$\frac{1}{3N}\cdot\frac{\partial \ln \overline{\omega^2}}{\partial V_{ij}}\cdot U_h$$

$$=\begin{cases}kT\dfrac{\partial}{\partial V_{ij}}\ln\overline{\omega^2} & \text{für hohe Temperaturen}\\ \dfrac{\hbar\bar{\omega}}{2}\dfrac{\partial}{\partial V_{ij}}\ln\overline{\omega^2} & \text{für } T=0\end{cases} \tag{A.8}$$

Zur Bildung der Mittelwerte können wir in erster Näherung das Debyesche Spektrum benutzen. Es ergibt sich (ω_v ist die Debyesche Abschneidefrequenz unter Einschluss der Anharmonizität) für hohe Temperaturen :

$$\frac{1}{2}\cdot\frac{\partial}{\partial V_{ij}}\overline{\ln\omega^2}=\frac{\partial}{\partial V_{ij}}(\ln\omega_v-0{,}333)=\frac{1}{\omega_v}\cdot\frac{\partial\omega_v}{\partial V_{ij}}$$

$$\frac{1}{2}\cdot\frac{\partial}{\partial V_{ij}}\ln\overline{\omega^2}=\frac{\partial}{\partial V_{ij}}(\ln\omega_v-0{,}255)=\frac{1}{\omega_v}\cdot\frac{\partial\omega_v}{\partial V_{ij}}$$

für $T=0$: (A.9)

$$\frac{1}{2}\cdot\frac{\partial}{\partial V_{ij}}\ln\overline{\omega^2}=\frac{\partial}{\partial V_{ij}}(\ln\omega_v-0{,}287)=\frac{1}{\omega_v}\cdot\frac{\partial\omega_v}{\partial V_{ij}}$$

$$\frac{1}{2}\cdot\frac{\partial}{\partial V_{ij}}\ln\overline{\omega^2}=\frac{\partial}{\partial V_{ij}}(\ln\omega_v-0{,}255)=\frac{1}{\omega_v}\cdot\frac{\partial\omega_v}{\partial V_{ij}}.$$

Hiernach sollte die Ersetzung exakt erlaubt sein! Wir dürfen uns aber nicht täuschen lassen. Beim richtigen Gitterspektrum wird auch der Verlauf des Spektrums von den mittleren Lagen A abhängen, nicht nur die Abschneidefrequenz. Das führt dazu, dass die Zahlen in (A. 9) durch Ausdrücke zu ersetzen sind, die von A abhängen. Immerhin sollte diese Abhängigkeit von A merklich kleiner sein als die der Abschneidefrequenz. Für Temperaturen zwischen $T=0$ und hohen Temperaturen dürfen wir annehmen, dass der Fehler von derselben Grössenordnung ist, da alle vorkommenden Grössen monotone Funktionen der Temperatur sind.

Wir wollen die Fehlerabschätzung noch von einer anderen Seite beleuchten. Es sei der Fall reiner Zentralkräfte mit Wechselwirkung nur zu nächsten Nachbarn betrachtet. Das Modell lässt sich durch eine effektive Federkonstante f beschreiben. Es gilt $\omega_s^2(\boldsymbol{k})=f\cdot\chi_s(\boldsymbol{k})$, wo $\chi_s(\boldsymbol{k})$ nicht von den Verzerrungen abhängt (vgl. Abschnitt 2), also $\ln\omega_s^2(\boldsymbol{k})=\ln f+\ln\chi_s(\boldsymbol{k})$ und $\partial\ln\omega_s^2(\boldsymbol{k})/\partial V_{ij}=\partial\ln f/\partial V_{ij}$, unabhängig von k! Die Ersetzung ist exakt erlaubt! Sobald wir aber Wechselwirkungen zwischen Nachbarn höherer Ordnung sowie Mehrkörperkräfte berücksichtigen müssen, liegen die Verhältnisse nicht mehr so einfach. Dann ist die Ersetzung sicher nicht mehr exakt. Den Hauptanteil werden aber auch in diesen Fällen die Zentralkräfte zwischen nächsten Nachbarn liefern, so dass der Fehler relativ klein sein wird. Alles in allem dürfen wir die gemachte Ersetzung von $\partial\ln\omega_s^2(\boldsymbol{k})/\partial V_{ij}$ durch $\partial\ln\overline{\omega^2}/\partial V_{ij}$ wohl als sehr gute Näherung ansehen.

LITERATUR

1. Leibfried G. *Handbuch der Physik* Bd. VII 1, S.104–324. Springer, Berlin (1955).
2. Born M. und Huang K. *Dynamical Theory of Crystal Lattices*. Oxford (1954).
3. Born M. *Nachr. Akad. Wiss. Göttingen* Heft 6 (1951).
   Born M. und Oppenheimer R. *Ann. Phys., Lpz.* **84,** 457 (1927).
4. Leibfried G. und Brenig W. *Fortschr. Phys.* **1,** 187 (1953); *Z. Phys.* **134,** 451 (1953).
5. Born M. und von Kármán Th. *Phys. Z.* **14,** 15, 65 (1913) vgl. auch Blackman M. *Handbuch der Physik* Bd. VII 1, S.325–382. Springer, Berlin (1955).
6. Debye P. *Ann. Phys., Lpz.* **39,** 789 (1912).
7. Leibfried G. und Schlömann E. *Nachr. Akad. Wiss. Göttingen* 71 (54).
8. Born M. *Nachr. Akad. Wiss. Göttingen, Festschrift* 1 (1951).
9. Hooton D. J. *Phil. Mag.* **46,** 422, 433, 485, 701 (1955); *Z. Phys.* **142,** 42 (1955).
10. Nakajima S. *Advanc. Phys.* **4,** 363 (1955).
11. Grüneisen E. *Ann. Phys., Lpz.* **26,** 393 (1908).
    Mie G. *Ann. Phys., Lpz.* **11,** 657 (1903).
12. Gow M. mit Born M. *Proc. Camb. Phil. Soc.* **40,** 151 (1944).
13. Born M. und Brody E. *Z. Phys.* **6,** 132 (1921).

VOL. 4 1958

14. SCHRÖDINGER E. *Z. Phys.* **11,** 170 (1922).
15. CLUSIUS K. *Z. Phys. Chem.* B**31,** 459 (1936).
16. QUIMBEY S. L. und SUTTON P. M. *Phys. Rev.* **91,** 1122 (1953).
17. HUSIMI K. *Proc. Phys.-math. Soc. Japan* **22,** 265 (1940).
18. LEIBFRIED G. und HAHN H. *Z. Phys.* (im Druck).
19. CLUSIUS K. *Z. Naturforsch.* **11a,** 715 (1956).
20. MARTIN D. L. *Phil. Mag.* **46,** 751 (1955).
21. DURAND M. A. *Phys. Rev.* **50,** 449 (1936). OVERTON W. C. *Phys. Rev.* **84,** 758 (1951).
22. RAMAN C. V. *Proc. Ind. Acad. Sci.* A**42,** 111 (1955); vgl. auch HEARMON R. F. S. *Advanc. Phys.* **5,** 323 (1956).

L.
58

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 306–314.

CRYSTALLINITY CHANGES IN IRRADIATED POLYETHYLENES

A. CHARLESBY and L. CALLAGHAN

Royal Military College of Science, Shrivenham, Wilts.
Tube Investment Research Laboratories, Hinxton Hall, Cambridge

(*Received* 12 *July* 1957)

Abstract—The destruction of crystallites in polyethylene exposed to radiation has been followed by measurements of specific volume. For equal energy absorption, pile radiation is far more effective than electrons in disrupting crystallites, possibly because of the interaction of fast neutrons with the hydrogen atoms. The reduction in crystal size can be readily followed by using the concept of equivalent paraffin size outlined in a previous paper. The larger crystals are more readily destroyed in both low- and high-density polyethylene.

An extrapolation of the specific-volume temperature curve to 0°K gives an approximately constant value v_0 for the specific volume, whether the material be amorphous, lightly crosslinked (as in a rubber), highly crosslinked (as in a glass-like state), or crystalline. The specific volume v at any temperature T can be written in the form

$$(v-v_0)(P+P_0+P_r r) = RT/M,$$

where P is the external pressure, P_0 the internal pressure, $P_r r$ is the equivalent of an internal pressure arising from the crosslinks produced by a radiation dose r, R is the gas constant and M the unit molecular weight (28 for C_2H_4 in polyethylene).

VOL. 4 1958

INTRODUCTION

In a previous paper,[1] a study was made of the crystallinity of samples of polyethylene having varying degrees of branching. An earlier paper by Charlesby and Ross[2] drew attention to the reduction in the maximum melting point of low-density (branched) polyethylene resulting from exposure to atomic-pile radiation at about 80°C. In the present paper, this effect is studied in greater detail and compared with the effect produced by an equal dose of electron radiation. Other forms of polyethylene with more extensive crystalline regions have also been investigated. The observed changes in crystallinity are expressed in terms of equivalent crystal size, as presented in the previous paper.[1]

ATOMIC-PILE RADIATION

Fig. 1 shows the specific-volume–temperature curves obtained for low-density polyethylene irradiated in the BEPO atomic pile over a very wide range of radiation doses. The measurements of specific volume were carried out under the same conditions as those described in the previous paper[1] and are therefore directly comparable with the data obtained for unirradiated polyethylene.

The curves show that with increasing exposure to radiation there is a pronounced decrease in the maximum melting temperature, i.e., the lowest temperature at which all crystals have melted. Above this temperature unirradiated polyethylene becomes a viscous liquid, since the van der Waals forces are no longer sufficient to bind the polymer molecules together in a regular array. In irradiated polyethylene, on the other hand, the disappearance of crystallinity still leaves radiation-induced crosslinks available to bind the molecules. The destruction of all crystals due to the combination of heat and radiation damage then gives rise to a rubber-like amorphous structure, whose properties depend primarily on the density of crosslinks.

The progressive decrease in the maximum

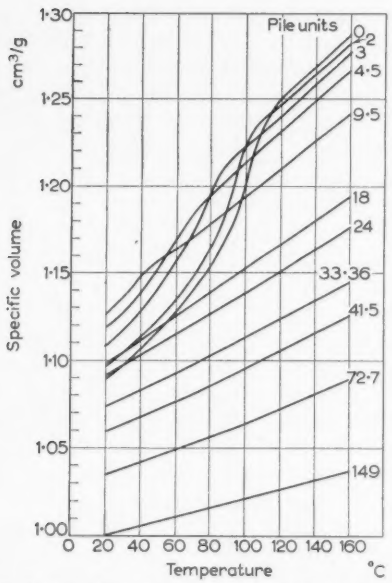


FIG. 1. Specific volume of irradiated polyethylene Alkathene 2. Pile irradiation; Figures show radiation dose in pile units.

crystalline melting point is ascribed to the preferential destruction of the larger crystals by radiation. For low and medium doses of radiation, some crystals remain which are stable at room temperature, but at the higher radiation doses all crystallinity has disappeared even at room temperature, leaving an amorphous material which can show either rubber-like or glass-like properties, depending on the density of crosslinks. Unirradiated low-density polyethylene, for example, shows a maximum crystalline melting point of 115°C; after a radiation dose of 4½ pile units, this figure is decreased to about 70°C, owing to the preferential destruction of the larger crystallites. At a radiation dose of 18 pile units, no crystals remain even at room temperature, and the properties of the polymer correspond to a highly crosslinked amorphous network. Small crystals may indeed reform at much lower temperatures, but owing to their limited size and low thermal stability, these melt below room temperature.

Specific volume of irradiated polyethylene above the melting point

At temperatures above the crystal melting point, when no crystals remain, the specific volume expands uniformly with temperature, although for irradiated polyethylene the coefficient of thermal expansion is lower than for unirradiated polymer. At any given temperature, the reduction in specific volume due to radiation arises primarily from the presence of crosslinks which bind the molecules more closely together. Loss of hydrogen due to radiation has the opposite effect on the specific volume, the weight of a specimen with a given number of carbon atoms being reduced; the correction for this loss of weight is small compared with that for the contraction due to crosslinking.

Crystallinity in irradiated polyethylene

To determine the degree of crystallinity present at various temperatures and radiation doses, the same procedure was used as for unirradiated materials.[1] No significant change is observed in the spacings of the X-ray diffraction pattern as a

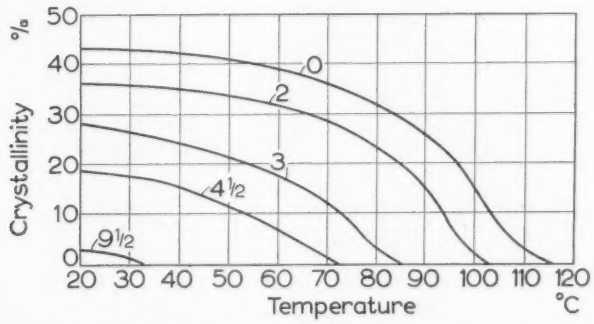


FIG. 2. Crystallinity of irradiated polyethylene Alkathene 2 dose in pile units Vc = 1·00; $\alpha = 4 \times 10^{-4}$.

result of radiation, so that the value chosen for the specific volume of the crystalline regions v_c is unaltered. However, crosslinks cause a contraction in the specific volume of the amorphous regions; the specific volume v_a at any temperature can be obtained only by a linear extrapolation from above the maximum crystal-melting temperature. The crystalline fractions deduced from the specific volumes are shown in Fig. 2 for a low-density polyethylene subjected to various radiation doses. From the variation of the specific volume with temperature, the equivalent destruction of n paraffins C_nH_{2n+2} can be derived with appropriate values for n (Fig. 3). On this view, the effect of

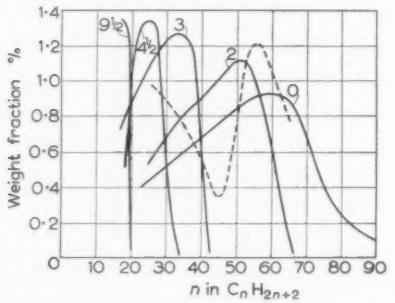


FIG. 3. Equivalent size distribution for irradiated alkathene 2. (Figures indicate radiation dose in pile units). Broken curve indicates changed distribution for electron radiation at 80°C 150 Mrad.

radiation is primarily to disintegrate individual crystals into smaller crystals, some of which may be too small to be stable even at room temperature. The apparent increase in the number of small crystals with increasing radiation dose would then arise from the partial disruption of the larger crystals giving smaller crystals, less the destruction of smaller crystals by radiation. From the figure, it would appear that a radiation dose of two pile units destroys all crystals of equivalent size greater than C_{65}.

As in the case of unirradiated polyethylene, the size of these crystallites may be limited by the density of branches and in addition by the presence of crosslinks which do not fit readily into the crystal structure. Table 1 shows that for any radiation dose the most frequent crystal size, expressed in equivalent linear paraffins C_nH_{2n+2} agrees approximately with that calculated from the average distance between branches or crosslinks. In this calculation, one pile unit of radiation is taken to produce one crosslinked unit per 200 carbon atoms, as found previously, while the initial branching density amounts to about one branch per 60 carbon atoms for the Alkathene 2 considered here.

ELECTRON RADIATION

Data on the specific volumes of similar polyethylene specimens subjected to varying doses of electron radiation are shown in Fig. 4. In this case, the high intensity of the electron beam allowed the radiation dose to be accumulated at the rate of 1 Mrad*/sec (equivalent to a dose rate of approximately one pile unit per min instead of per 16 hr,

* 1 Mrad is the dose corresponding to an energy absorption of 10 J/g.

Table 1. Comparison of average distance between successive branches or crosslinks and equivalent crystal size

Dose (pile units)	Branches* per 100 C atoms	Units crosslinked per 100 C atoms	Total per 100 C atoms	Average number of C atoms between branches or crosslinks	Peak Frequency† observed (n in C_nH_{2+n})
0	1·67	0	1·67	60	60
2	1·67	1	2·67	37·5	50
3	1·67	1·5	3·17	31·5	32·5
4½	1·67	2·25	3·92	25·5	25
9	1·67	4·5	6·17	16·2	17

* Assuming no change with radiation.
† From equivalent n paraffin distribution curve in Fig. 3.

VOL. 4 1958

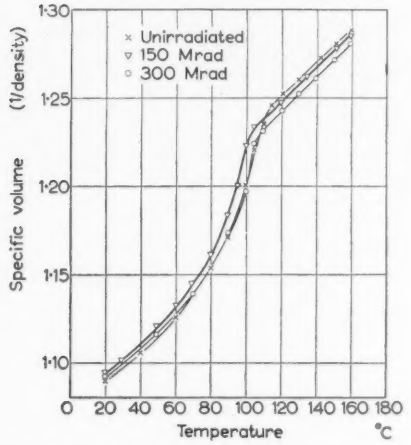


FIG. 4. Volume of electron irradiated Alkathene 7.

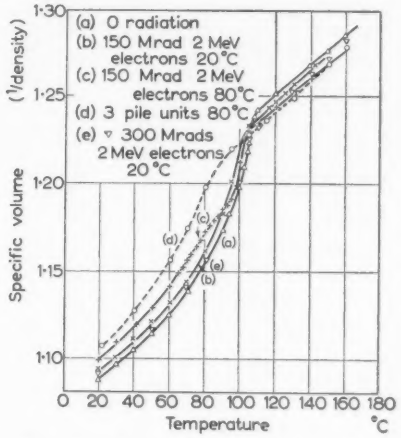


FIG. 5. Specific volume of polyethylene (Alkathene 7).

as in the atomic pile). To keep the temperature at about 20°C, suitable time intervals were allowed between successive doses. The general shape of the specific-volume temperature curve is similar to that obtained for atomic-pile irradiation, but the maximum temperature for crystal melting is not lowered to the same extent as for pile radiation of equal energy absorption. For a dose of 300 Mrad, this decrease amounts to only about 10°C, whereas for the equivalent pile dose of 6 units, the reduction would be 55°C. Thus, electron irradiation at 20°C has a much smaller effect in destroying crystals than has pile radiation at about 80°C.

This difference is not to be ascribed to the increased temperature in the pile, which would reduce the stability of crystals, nor to the increased effectiveness of radiation in promoting crosslinking at higher temperatures. Curves in Fig. 5 compare the effect of temperature of irradiation (20 and 80°C in each case for 150 Mrad of electron irradiation) with that for 3 units of pile radiation (equivalent to 150 Mrad or 1500 Jg of energy absorption). It is seen (curve c) that for electron radiation at 80°C two transition temperatures appear. The higher is not appreciably affected by changes in the temperature of radiation, whereas the lower falls at about the same position as that for pile-irradiated polyethylene. These two transitions are made clearer in Fig. 3.

It may therefore be assumed that the differences arise from the type of radiation used. This conclusion may at first appear surprising, since most experiments on radiation effects in organic materials indicate little difference between lightly ionizing radiation, such as fast electrons, X-rays, or γ-rays, and highly ionizing radiations, such as fast protons or α-particles. A possible explanation may be advanced in terms of the particles involved. When fast electron radiation is used, the individual ions produced are scattered approximately at random throughout the polymer, whereas with pile radiation fast neutrons contribute part of the ionizing effect by colliding with hydrogen atoms. The fast protons produced in this way have a very high ionizing power with a correspondingly short range. In the vicinity of each primary collision site, there will be a considerable evolution of kinetic energy, with resultant dislocations somewhat similar to those occurring in irradiated metals. The probability of a region being formed in the crystal with a large number of adjacent dislocations is therefore greatly enhanced when atomic-pile radiation containing fast neutrons is used. Such a region of intense local disorder will have a higher probability of disrupting a crystal than an equal number of dislocations distributed at random, and in a crystalline polymer will cause more permanent damage than in a metal.

For electron radiation at about 80°C, it appears that crystals of equivalent size to between $C_{40}H_{82}$ and $C_{50}H_{102}$ (melting points 82–92°C) are particularly susceptible to destruction by radiation.

Smaller crystal sizes would melt under the radiation temperature and reform quite readily on subsequent cooling. The largest crystals, on the other hand, persist at the temperature of radiation, and the low density of ionization of electrons is insufficient to cause them to disrupt.

LINEAR POLYETHYLENE

The change in specific volume with radiation dose for the more highly crystalline low-pressure polyethylene is shown in Figs. 6 and 7. As for high-pressure polyethylene, the effect of doses of 300–350 Mrad on the transition temperature is small, being of the order of 5–10°C. Above this

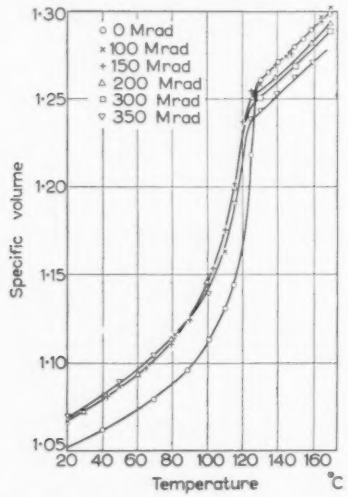


FIG. 6. Specific volume of polyethylene (Ziegler) (2 MeV electrons).

transition temperature, the specific volume of the amorphous regions is reduced by radiation owing to the formation of crosslinks. At 150°C, for example, it decreases from 1·283 for Marlex and 1·280 for Ziegler to 1·270 or 1·269 after a radiation dose of 300 Mrad. The changes in Ziegler material do not appear to be linear with radiation dose, although as these changes are relatively small, this non-linearity may arise from small experimental errors.

DENSITY OF CROSSLINKED AMORPHOUS REGIONS

Measurements of the solubility, swelling, and

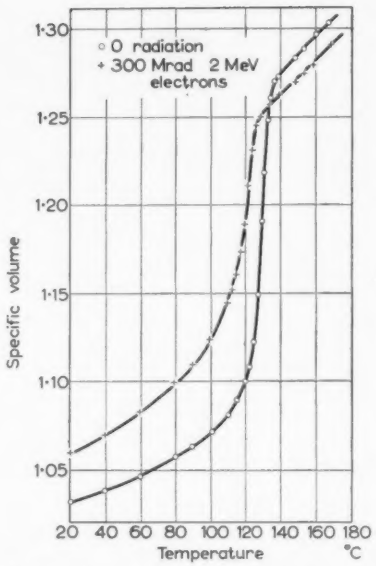


FIG. 7. Specific volume of polyethylene (Marlex 50).

elastic modulus of polyethylene have shown that the degree of crosslinking for a given dose depends on the irradiation temperature, increasing with the latter.[3,4] On the other hand, differences in the degree of crystallinity (as in different types of polyethylene) appear to have little effect on the efficiency of crosslinking. These effects should also appear when the contraction in volume due to the formation of C–C bonds between molecules is considered. For such comparisons it is necessary to consider the specific volume above the crystal melting point, when changes in the percentage of crystallinity are no longer a complicating factor.

Table 2 shows the decrease in specific volume at 150°C for a number of polyethylene samples, subjected to a small dose (100 Mrads or 2 pile units) or to higher doses of electron or pile radiation. At 20°C radiation reduces the amorphous volume by $0{\cdot}4\times10^{-2}$, whereas when irradiation is carried out at 80°C, the reduction is considerably greater (about 1×10^{-2}). This indicates the increased efficiency of radiation-induced crosslinking at the higher temperature. No significant difference was observed as between Alkathene Grade 7 and Ziegler or Marlex material, although there is a difference (by a factor of about three) in the amount of amorphous material present,

VOL. 4 1958

Table 2. Initial decrease in specific volume v *at* 150°C *per* 100 Mrad (*or 2 pile units*)

Grade of polyethylene	Type*	Radiation	Irradiation temperature (°C)	Decrease v_{150}	Dose Range up to:	ΔV_{150} (volume change per 100 Mrad)
Alkathene 7	HP	Electron	80	$1{\cdot}7\times10^{-2}$	150 Mrad	$1{\cdot}1\times10^{-2}$
Alkathene 7	HP	Electron	20	$1{\cdot}2\times10^{-2}$	300 Mrad	$0{\cdot}4\times10^{-2}$
Alkathene 7	HP	Pile	80	5×10^{-2}	10 pile units	$1{\cdot}0\times10^{-2}$
Ziegler	LP	Electron	20	$1{\cdot}1\times10^{-2}$	300 Mrad	$0{\cdot}37\times10^{-2}$
Marlex	LP	Electron	20	$1{\cdot}2\times10^{-2}$	300 Mrad	$0{\cdot}4\times10^{-2}$
Alkathene 2	HP	Pile	80	12×10^{-2}	24 pile units	$1{\cdot}0\times10^{-2}$

* HP = high pressure, LP = low pressure.

confirming that the effect does not depend on the degree of crystallinity.

Changes in specific volume of the amorphous region due to radiation effects arise from the evolution of hydrogen, the formation of crosslinks, and the increase in unsaturation. The loss of hydrogen can be estimated from the change in the H/C ratio, since almost all the gas evolved by radiation is hydrogen. Microchemical analysis of irradiated polyethylene leads to the results shown in Fig. 8. There appears to be a rapid drop at radiation doses below about 10 pile units, followed by a less rapid decrease up to about 100 pile units. The initial slope amounts to the loss of 1 H atom per 100 C atoms per pile unit ($C_{100}H_{200}$ becomes $C_{100}H_{190}$ per 10 pile units); the corresponding weight loss is 0·7 mg/g of polyethylene per pile unit, and the G value for hydrogen evolution $G(H_2) \cong 7$. This loss of hydrogen will arise both from the increased unsaturation and from the formation of crosslinks. Above about 10 units, unsaturation tends to a limiting value and the reduced loss of hydrogen amounting to only 0·4 H atom per 100 C atoms per pile unit, would arise only from the formation of crosslinks. Then $G(H_2) \cong 3$. This value is in approximate agreement with estimates for crosslinking deduced by other means. The contraction in volume of the amorphous region due to crosslinking can be calculated from these data. These estimates are very

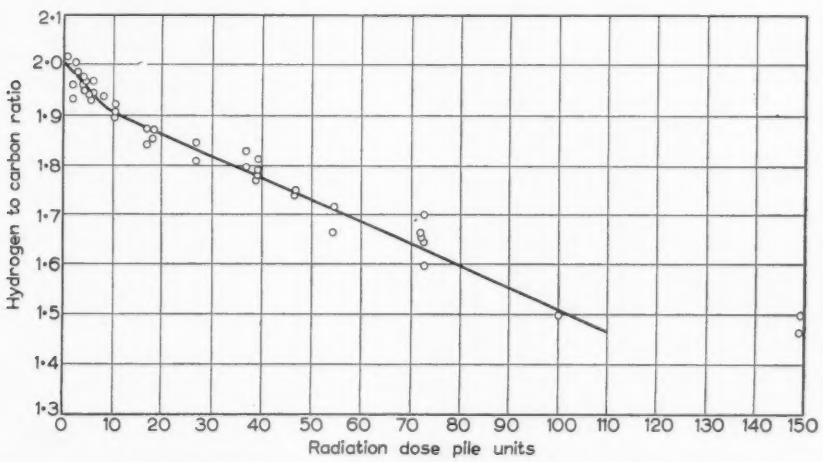


FIG. 8. Hydrogen to carbon ratio in irradiated polyethylene.

approximate; a more accurate method is to plot the specific volume (at 150°C) directly as a function of the H/C ratio (Fig. 9). This figure also shows the volume occupied by each carbon atom (plus the associated hydrogens) at 150°C, allowing for the loss in weight due to the hydrogen atoms evolved.

With the possible exception of a transition for highly irradiated polyethylene at about 70°C, the specific volume of amorphous material is linear with temperature over the range studied. By assuming that the uniform expansion extends to low temperatures, a value can be given for the

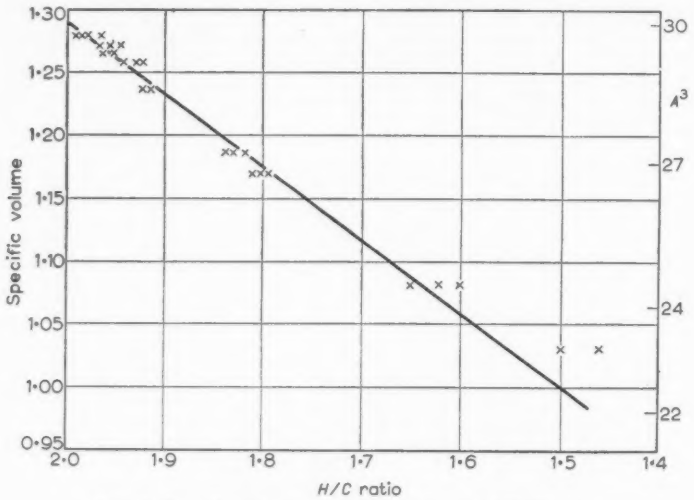


FIG. 9. Specific volume at 150°C in terms of H/C ratio.

It is seen that a H/C ratio of 1·5, corresponding to one crosslinked unit every two carbon atoms, causes a reduction in volume per carbon atom from 30 to 22·5 Å³. Thus, each crosslink, involving the loss of two hydrogen atoms from two carbon atoms, contracts the volume by about 30 Å³. Since the total volume occupied by two (CH_2) groups averages 60 Å³, it follows that the maximum degree of crosslinking will in any case be limited by steric considerations. This appears to be the case at very high doses, of the order of 150 pile units, when the contraction in the specific volume tends to approach a limit, defined by the minimum volume which can be occupied by carbon atoms bound to each other by primary bonds. Steric factors will also limit the density of crosslinks possible in a polyethylene structure. For a crosslink between two molecules to be formed, it is necessary for them to lie at a steep angle, and not parallel with one another as in a crystal. The destruction of crystallinity may therefore be ascribed to the rearrangement necessitated by the introduction of a crosslink.

specific volume v_0 at 0°K as a function of the degree of crosslinking, in the absence of thermal vibration. The values obtained are shown in Table 3, which also gives the specific volume of the crystalline region, extrapolated from a value of 1·00 at 0°C by assuming a uniform coefficient of thermal expansion of 4×10^{-4} at 0°C.

Table 3. Specific volume at 0°K (v_0)

Radiation dose (pile units)	v_0(cm³/g)	$\alpha^* \times 10^3$
0	0·87	1·11
4·5	0·877	1·025
9·5	0·898	0·93
18	0·896	0·765
41·5	0·906	0·56
72·7	0·917	0·435
149	0·924	0·28
Crystalline	0·891	0·45

* α is the change in volume per °C, divided by v_0 at 0°K.

VOL. 4 1958

The values of v_0 show surprisingly little difference, although they are obtained from polyethylene in very different states: liquid (0 dose), rubber-like (4·5 units), glass-like (18 units and above), and crystalline. It is therefore possible to express the specific volume in the range of temperatures considered here in the approximate form

$$v(T) = 0{\cdot}88(1+\alpha T)$$

where only α depends on the state of the polymer-crosslinked, crystalline, or amorphous.

SPENCER and GILMORE[5] have measured the specific volume v of several plastics at various external pressures. Their results can be expressed in the form

$$(v-v_0)(P+P_i) = RT/M$$

where v_0 is a parameter, P the external pressure, P_i a parameter equivalent to an internal pressure, R is the gas constant, and T the absolute temperature. M is taken as the molecular weight (28) of ethylene; v_0 is the volume (deduced by extrapolation) at absolute zero. The value of v_0 derived by them is 0·875, in excellent agreement with our own value.

The formula can be written:

$$v = v_0+\frac{RT}{(P+P_i)M}$$

so that with our data $\alpha = R/(P+P_i)Mv_0$. Since the external pressure P in our experiments was atmospheric, and therefore negligible compared with P_i, we may write (with $v_0 = 0{\cdot}88$)

$$\alpha = R/0{\cdot}88P_iM.$$

The decrease in the volume coefficient α due to crystallinity or crosslinking can therefore be considered as equivalent to an increased internal pressure P_i or alternatively as an increase in the molecular weight M, consistent with the decreased number of degrees of freedom per (CH_2) group. The difference between these two explanations can be decided by measurements of the specific volume of irradiated polyethylene under pressure.

The value derived for P_i by SPENCER and GILMORE[5] was $3{\cdot}3\times10^9$, so that $R/0{\cdot}88\ P_iM = 1{\cdot}02\times10^{-3}$, in approximate agreement with the experimental value of α for unirradiated amorphous polyethylene (Table 3). If the effect of linking molecules together is equivalent to raising the internal pressure by an amount proportional to the number of crosslinks or the radiation dose r,

$$P_i = P_0+P_rr,$$

where P_r is the increased internal pressure per unit radiation dose, then

$$\frac{1}{\alpha} = \frac{0{\cdot}88P_iM}{R} = 3\times10^{-7}[P_0+P_rr].$$

A plot of $1/\alpha$ against r does give the required linear plot (Fig. 10) with values for P_0 and P_r of about

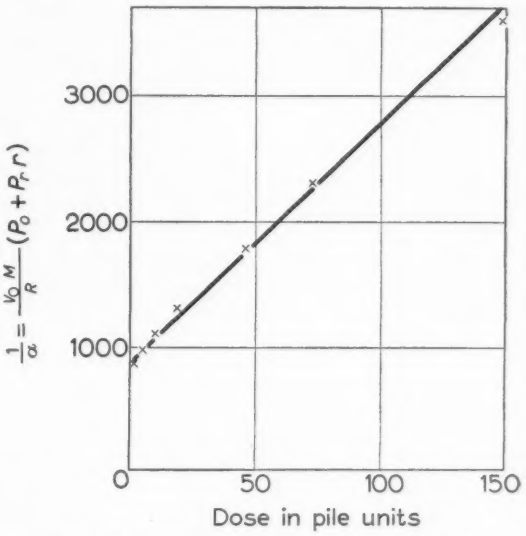


FIG. 10. Expansion of irradiated polyethylene $1/\alpha = 3\times10^{-7}\ (P_0+P_rr)$.

3×10^9 and $6{\cdot}4\times10^7$. Thus, a radiation dose of about 50 pile units, or 25 per cent crosslinking, is needed to double the internal pressure. Crystallinity also increases the internal pressure from 3×10^9, for amorphous material, to $7{\cdot}4\times10^9$ dynes/cm^2.

CONCLUSIONS

Exposure of polyethylene to high-energy radiations from an atomic pile results in the destruction of crystalline regions. The larger crystallites are most affected, so that the maximum melting point is decreased. For a radiation dose of about 10 units

(corresponding to an energy absorption of about 5×10^3 J/g), no crystals stable at room temperature remain. The destruction of crystals can be readily traced by using the concept of equivalent crystal size described in a previous paper.[1]

Electron radiation appears to be less effective than atomic-pile radiation in causing disruption of crystallites. The difference may arise from the intense local dissipation of energy during the collision of a fast neutron (present in atomic piles) with a proton or carbon atom. The largest crystals initially present in the unirradiated material appear to be more resistant to electron radiation, where the energy dissipation is less localized.

In spite of the considerable difference in crystal dimensions, the effect of electron radiation on the crystallinity of other forms of polyethylene follows the same pattern as that observed for low-density polyethylene.

The specific volume of various forms of polyethylene–crystalline, amorphous, or crosslinked when extrapolated back to 0°K gives an approximately constant value v_0 of about 0·88. The specific volume v observed at higher temperatures can be expressed in terms of different coefficients of thermal expansion α, or of different external or internal pressures P or P_i due to differences of binding between molecules. For a crosslinked polyethylene, the internal pressure P_i is proportional to the density of crosslinking q:

$$P_i = 3\times10^9+11\times10^9q,$$

The value 3×10^9 is in approximate agreement with the value obtained for the internal pressure of unirradiated polyethylene by other methods. The specific volume v of crystalline, amorphous, or crosslinked polyethylene can therefore be expressed by a simple formula

$$(V-V_0)(P+P_i) = RT/M,$$

which is somewhat similar in form to that representing the behaviour of gases.

VOL. 4 1958

REFERENCES

1. Charlesby A. and Callaghan L. *J. Phys. Chem. Solids* **4,** 227 (1958).
2. Charlesby A. and Ross M. *Proc. Roy. Soc.* A **217,** 122 (1953).
3. Black R. M. *Nature, Lond.* **178,** 305 (1956).
4. Charlesby A. and Davison W. H. T. *Chem. Ind. (Rev.)* p. 232 (1957).
5. Spencer R. S. and Gilmore G. D. *J. Appl. Phys.* **20,** 502 (1949); **21,** 523 (1950).

J. Phys. Chem. Solids Pergamon Press 1958. Vol. 4. pp. 315–317.

ABSORPTION SPECTRUM OF BISMUTH-DOPED SILICON

H. J. HROSTOWSKI and R. H. KAISER

Bell Telephone Laboratories, Inc., Murray Hill, N.J.

(*Received* 29 *August* 1957)

Abstract—The infrared absorption spectrum of bismuth-doped silicon at 4·2°K indicates a more detailed donor-term scheme than has previously been observed for other Group-V impurities. Positions of *p*-type levels relative to the conduction band are in excellent agreement with existing theory. A "forbidden" transition to the degenerate 2*s* levels is also observed. The optical ionization energy is $70{\cdot}6\times10^{-3}$ eV, in good agreement with the thermal value of 69×10^{-3} eV.

INTRODUCTION

ELECTRONIC spectra of phosphorus, arsenic, and antimony in silicon have been reported by PICUS *et al.*[1] Their results indicate that some excited states of neutral Group V donors are well described[1,2] by the effective-mass theory of KOHN and LUTTINGER.[3] However, the observed spectra show less similarity than is to be expected from theory. The variations may arise in part from the presence of several donor impurities with overlapping spectra.

A value 0·069 eV has recently been determined[4] for the thermal ionization energy of bismuth in silicon. Absorption of this donor occurs in a spectral region free from interference of other Group-V impurities. The spectrum of bismuth-doped silicon shows several interesting features not observed in other donor spectra, but in agreement with existing theory.

EXPERIMENTAL RESULTS

Transmission spectra were taken with a Perkin-Elmer Model 112 spectrometer with CsBr optics. The experimental methods are described elsewhere.[5] The results shown in Fig. 1 were obtained at 4·2°K with samples from a pulled crystal with an approximate room-temperature resistivity of 0·4 Ωcm and a net donor density of $2\times10^{16}/\text{cm}^3$. Sample thickness and spectral resolution are also given in Fig. 1. Table 1 lists the energies of the observed transitions. Except where otherwise noted, experimental error is less than 4×10^{-5}eV.

Table 1. *Assignments of observed transitions*

Transition	Assignment	Energy (eV $\times 10^3$)	Relative intensity	
			Calculated[2]	Observed*
I	$1s \rightarrow 2p,\ m = 0$	59·50	0·4	s
II	$1s(1) \rightarrow 2s(5)$†	62·19	—	vw
III	$1s \rightarrow 2p,\ m = \pm1$	64·57	1·0	vs
IV	$1s \rightarrow 3p,\ m = 0$	65·47	0·04	w
V	$1s \rightarrow$?	67·13	—	w
VI	$1s \rightarrow 3p,\ m = \pm1$	68·01	0·3	m
VII	$1s \rightarrow$?	$69{\cdot}12\pm0{\cdot}12$	—	w

* vs = very strong, s = strong, m = moderate, w = weak, vw = very weak.
† The parenthetical numbers give the degeneracy of these states.

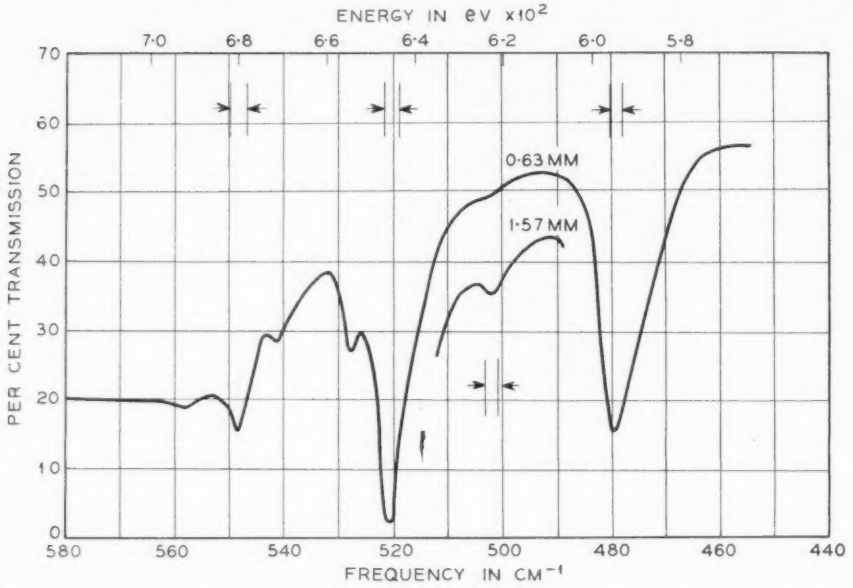


FIG. 1. The electronic spectrum of bismuth in silicon at 4·2°K. Approximate effective slit widths are as shown.

Tentative assignments, obtained from comparison of the observed and calculated relative intensities, are also included in Table 1. Two weak absorptions of pulled silicon crystals also occur between 580 and 500 cm^{-1}. However, all the bands shown in Fig. 1 decreased in intensity on heating. This behavior is characteristic of electronic transitions and opposite to that expected for lattice vibrations.

DISCUSSION

The major difficulty in the effective-mass formulism is a lack of detailed knowledge of the true potential in the immediate vicinity of an impurity center. Thus, observed ionization energies vary from the calculated energy by as much as a factor of three. The energies of higher s-states are also imperfectly known.

However, p-states* are affected very little by the specific nature of the donor atom. The positions relative to the conduction band of the $(2p, m = 0)$, $(2p, m = \pm 1)$, and $(3p, m = \pm 1)$ states are essentially constant and in excellent agreement with theory for phosphorus, arsenic, and antimony.[1] This is also true for bismuth, as shown in Table 2, which lists the calculated and observed spacings of the excited states of bismuth. Since we have resolved more transitions than have been observed for other donors, a more detailed comparison of theory and experiment is possible. The following observations are noteworthy:

(1) Theory predicts that the $(3p, m = 0)$ state should be much less intense than the $(2p, m = \pm 1)$ state, and should occur at a slightly higher energy. The transition from the ground state to the $(3p, m = 0)$ level, which has not been resolved for phosphorus, arsenic, or antimony, is clearly resolved for bismuth, as shown in Fig. 1. Its observed separation from the $(2p, m = \pm 1)$ state agrees with calculated value[2] within the error of the calculation.

(2) The very weak band observed at 501·7 cm^{-1} is definitely not a lattice vibration or an absorption of any known donor. We assign this as a transition from the non-degenerate ground state to five-fold degenerate $2s$-state. To a first approximation, this transition is forbidden. However this $2s$-state is only approximately five-fold degenerate and consists of two strictly two-fold and three-fold

* We use the notation of KOHN and LUTTINGER.[3]

VOL. 4 1958

Table 2. Spacings of excited states of bismuth

States	Energy difference ($eV \times 10^3$)	
	Theory	Experiment
$(2p, m = \pm 1) - (2p, m = 0)$	$5{\cdot}0 \pm 0{\cdot}3$	$5{\cdot}07 \pm 0{\cdot}06$
$(3p, m = 0) - (2p, m = \pm 1)$	$0{\cdot}2 \pm 0{\cdot}7$	$0{\cdot}905 \pm 0{\cdot}075$
$(3p, m = \pm 1) - (2p, m = \pm 1)$	$3{\cdot}0 \pm 0{\cdot}15$	$3{\cdot}45 \pm 0{\cdot}075$
$[2s(5)] - (2p, m = 0)$	$1{\cdot}1 \pm 0{\cdot}3$*	$2{\cdot}69 \pm 0{\cdot}06$
? $(2p, m = \pm 1)$	Not calculated	$2{\cdot}565 \pm 0{\cdot}075$
? $(3p, m = 0)$	Not calculated	$1{\cdot}10 \pm 0{\cdot}16$

* The theoretical position of the "five-fold" degenerate 2*s*-state was estimated from the considerations given in ref. (3).

degenerate sets of levels.[3] Observation of this transition indicates that the $2s(5)$ levels have some *p*-character which increases the probability of a transition from the totally *s*-type ground state. This assignment accounts for the very low intensity of the band which is characteristic of "forbidden" transitions. The agreement between the observed and calculated separation from the $(2p, m = 0)$ level is reasonable, since the nature and energy of the $2s(5)$ level can only be roughly estimated from the effective-mass theory.

(3) Two other absorptions are observed at 541·6 and 557·6 cm^{-1}. In the absence of detailed theoretical results, no reliable assignments can be made.

(4) The optical ionization energy, E_0, can be obtained from the observed energies of transitions I, III, and VI (see Table 1) and the calculated separations of the corresponding *p*-states from the conduction band.[2] A value of $E_0 = 70{\cdot}6 \pm \pm 0{\cdot}3 \times 10^{-3}$ eV results. This is in good agreement with the thermal value 69×10^{-3} eV.[4]

Although the ionization energy of bismuth is 50 per cent higher than that of other Group-V donors, their energy-level schemes are all very similar. The experimental results for *p*-states are in excellent agreement with the effective-mass theory. Observation of the "forbidden" transition from the ground state to the $2s(5)$ levels of bismuth may be helpful in determining the true potential in the proximity of the donor atom.

Acknowledgement—We are grateful to Miss ARLENE MILLS for providing the crystal used for this work.

REFERENCES

1. PICUS G., BURSTEIN E. and HENVIS B. *J. Phys. Chem. Solids* **1,** 75 (1956).
2. KOHN W. *Phys. Rev.* **98,** 1856 (1955).
3. KOHN W. and LUTTINGER J. M. *Phys. Rev.* **98,** 915 (1955).
4. MORIN F. J. Private communication.
5. HROSTOWSKI H. J. and KAISER R. H. *Phys. Rev.* In press.

Superconductivity in the Pd–Se system

(*Received* 4 *September* 1957)

A FEW years ago, the rhodium–sulfur system was explored for superconductivity and a superconductor with the probable formula Rh_9S_8 was found.[1] This phase is cubic, belongs to one of the three space groups $O_h(1)$—$Pm3m$, $O(1)$—$P432$ or $T_d(1)$—$P\bar{4}3m$, and has a lattice constant $a_0 = 9{\cdot}911 \pm 0{\cdot}001$ Å.* It seems necessary that the structure determination decides the correct formula of the compound; density measurements are inadequate for this purpose. It is hoped that we shall be able to discuss this more fully at some future date. Also reported in the above-mentioned paper is the investigation of several other systems which contain superconductors. These include the Rh–Se system and the Ni–Pd–As system (see also references 2 and 3). At the time, the 'compound PdSe' was also checked for superconductivity, but the conductivity was found to be normal down to 1·0°K. The powder photograph of the 'PdSe' preparation leaves little doubt that it is isostructural with Rh_9S_8; this has recently been discussed by GRØNVOLD and RØST[4] in a paper on the Pd–S, Pd–Se, and Pd–Te systems. GRØNVOLD and RØST point out that the density measurements of $Pd_{1\cdot1}Se$ indicate a possible unit cell content of 34 Pd and 30 Se atoms. The space groups referred to above would allow this formulation.

Not only are the structural aspects of the Pd–Se system not yet entirely clear, but also the observations of superconductivity are not entirely explained. The appearance of the paper by GRØNVOLD and RØST[4] renewed our interest in the Pd–Se system. It has been known for some time that changes in stoichiometry (without change in crystal structure) may produce changes in superconducting transition temperatures. Thus melts were made with increasing Pd content. The results of these measurements are shown in Table 1.

We have found that in the palladium-rich preparations, it is possible for the '$Pd_{1\cdot1}Se$' phase (as designated by GRØNVOLD and RØST) to segregate from the more palladium-rich phase '$Pd_{2\cdot8}Se$'. Thus a powder photograph of such a preparation might not show the $Pd_{2\cdot8}Se$ phase. Visual examination of the sample indicates that the latter phase is usually duller than the $Pd_{1\cdot1}Se$ phase and does usually occlude some of the $Pd_{1\cdot1}Se$ phase.

The data in Table 1 indicate that as the palladium content increases, the $Pd_{1\cdot1}Se$ phase becomes superconducting. But there is no obvious change in the powder photograph of the $Pd_{1\cdot2}Se$ preparation with respect to either intensity or spacing. This situation may be analogous to that which occurs in the Pd–As system.[3,5] The observation that not all

* Some analysis of Rh_9S_8 single-crystal X-ray data has been made at these Laboratories since the publication of the paper by MATTHIAS *et al.*,[1] but this work has not yet been completed.

Table 1 *Observations of superconductivity in Pd–Se system*

Composition of preparation (Pd : Se = x : 1)	Transition temp. (°K)	Percentage of material superconducting	Phases appearing in powder photo*
x			
1·0	< 1·0	0	A
1·1	< 1·0	0	A
1·2	2·5	50	A
1·3	2·5	75	A+B (faint)
1·5	2·5	100	A+B (faint)
1·7	2·5	100	A+B
2·0	2·2	100	A+B
2·5	2·3	100	A (faint)+B

* A ≡ $Pd_{1\cdot1}Se$; B ≡ $Pd_{2\cdot8}Se$.

VOL. 4 1958

of the $Pd_{1.2}Se$ is superconducting implies inhomogeneity of the sample, but this also does not show up in the powder photograph (as line broadening for example). With further increase in palladium content, the $Pd_{2.8}Se$ phase appears and at the composition $Pd_{1.5}Se$ the sample is completely superconducting.

It appears to us that the most reasonable explanation of the observations is as follows: The $Pd_{1.1}Se$ phase becomes superconducting at a composition other than that of the ideal formula. In addition, the $Pd_{2.8}Se$ phase is superconducting and probably has a higher superconducting transition temperature when its composition differs from that of the ideal formula of that phase. Also the transition temperature of the $Pd_{1.1}Se$ phase is optimum for a range of composition then begins to drop with increasing palladium content.

The whole Pd–Se system has not as yet been investigated for superconductivity; we hope to do this sometime in the future.*

* *Note added in proof:* Since this letter was submitted, time permitted the completion of our intended measurements. The phases Pd_4Se and $PdSe_2$ reported by GRØNVOLD and RØST are not superconducting above 1·5°K.

Acknowledgements—The authors wish to thank R. GOLDSTEIN for measuring the Rh_9S_8 photograph and V. B. BALA and C. E. MILLER for taking the photographs of the Pd–Se preparations.

B. T. MATTHIAS
S. GELLER

Bell Telephone Laboratories Inc.,
Murray Hill, N.J.

REFERENCES

1. MATTHIAS B. T. CORENZWIT E. and MILLER C. E. *Phys. Rev.* **93,** 1415 (1954).
2. GELLER S. and CETLIN B. B. *Acta Cryst.* **8,** 272 (1955).
3. GELLER S. and MATTHIAS B. T. *J. Phys. Chem. Solids* **4,** 156 (1958).
4. GRØNVOLD F. and RØST E. *Acta Chem. Scand.* **10,** 1620 (1956).
5. THOMASSEN L. *Z. Phys. Chem.* B **4,** 277 (1929).

BOOK REVIEW

J. FRIEDEL: **Les Dislocations.** Gauthier-Villars, Paris, 1956. 314 pp., $10.20 unbound, $11.05 bound.

THIS book is developed from a course of lectures on dislocations. The author disclaims originality but admits that in parts the presentation is slightly personal. The book is directed mainly towards engineers, especially those interested in the fundamentals of mechanical properties, but will prove invaluable to all interested in the subject. Part I deals with the properties of single dislocations, both stationary and in motion; the concepts of extended and of imperfect dislocations, jogs and climb, and equilibrium between vacancies, interstitials and dislocations are explained. Part II is a useful survey of information on the properties of dislocation arrays, and current ideas on work hardening. This is followed by a discussion of recovery processes and of flow and fracture in metallic crystals. Part III deals with the interaction of dislocations with other kinds of defects; impurities, precipitates, magnetic domain walls are treated so that a coherent account of yield points in many types of material is possible. There is a good bibliography which forms an excellent guide to the rather dispersed literature of the subject.

The presentation is everywhere down-to-earth and leaves no doubt of the significance, even to the practical metallurgical engineer, of the dislocation concept. From the point of view of those who wish to study the subject more closely, and perhaps do some research themselves into the mechanical properties of crystals, the book can perhaps be criticized as being too complacent. For example, there is hardly adequate indication that the controversy concerning the maximum speed of dislocations in real lattices is still alive, nor that the magnitude and significance of the Peierls–Nabarro force is hotly debated. All but the exceptionally critical reader are in danger of being overimpressed by the wide range of success of the present theories; a clearer indication of the growing points of the theory and of some of the extant difficulties and differences of opinion would have improved the balance.

The book is clearly and accurately printed on paper which is pleasant to handle. Readers unaccustomed to uncut pages may find them irritating in use. The index is not very full, but the detailed contents list largely compensates this blemish. All in all, it is a book likely to be valued by most workers in this field as a coherent statement of the importance of dislocation theory as a unifying principle to discuss mechanical properties of crystals. Its shortcomings are likely to trouble only the specialists who may find no detail at just those points in which they are most interested.

W. M. LOMER

VOL. 4 1958

L.
58